Synthesis of pyrimidine C-nucleoside analogues and triphosphate derivatives

Author: Heng Ming Chan

Boston College

The Graduate School of Arts and Sciences

Department of Chemistry

SYNTHESIS OF PYRIMIDINE *C*-NUCLEOSIDE ANALOGUES AND TRIPHOSPHATE DERIVATIVES

a thesis

by

HENG MING CHAN

submitted in partial fulfillment of the requirements

for the degree of

Master of Science

May 2008

ACKNOWLEGEMENTS

Foremost let me express my gratitude to Professor Larry W. McLaughlin for the support and guidance he has provided unconditionally for the past three years.

I have had the pleasure of working with numerous excellent postdoctoral fellows, graduate, and undergraduate students. I am indebted to Dr. Meena, Dr. Zhenhua Sun, and Dr. Allen Horhota for their eagerness to teach me new techniques and assistance throughout the experiments. I also would like to extend my gratitude to Dr. Han Chen and Dr. Nicholas Greco for sharing their expertise in various areas and to William Issa for providing the heterocycle compounds for the coupling reactions.

Finally, I thank my family and friends both inside and outside of the laboratory for their moral support. I have grown stronger and more independent because of them.

TABLE OF CONTENTS

TABLE OF FIGURES

INTRODUCTION

RESULTS AND DISCUSSION

TABLE OF SCHEMES

RESULTS AND DISCUSSION

TABLE OF ABBREVIATIONS

3TC	(-)-2'-deoxy-3'-thiacytidine
A	2'-deoxyadenine
ABC	Abacavir
AIDS	acquired immunodeficiency syndrome
AMV-RT	Avain Mylelobastosis Virus reverse transcriptase
ATP	adenosine 5'-triphosphate
AZT	2',3'-dideoxy-3'-azidothymidine
C	2'-deoxycytidine
cDNA	chromosomal 2'-deoxyribonucleic acid
d	doublet
d*C	2-amino-5-(2-deoxy-β-D-ribofuranosyl)-pyridine
d*T	3-methylpyrin-2(*1H*)-one
d2APy	same as d*C
D4T	stavudine
dba	dibenzylideneacetone
dC	2'-deoxycytidine
DCM	dichloromethane
dd2APy	2-amino-5-(2,3-dideoxy-β-D-ribofuranosyl)-pyridine
dd2APyTP	2-amino-5-(2,3-dideoxy-β-D-ribofuranosyl)-pyridine 5'-triphosphate
ddC	2',3'-dideoxycytidine
ddI	2',3'-dideoxyinosine
ddNTP	2',3'-dideoxynucleoside 5'-triphosphate
DMAP	*N*,*N*-dimethylaminopyridine
DMF	*N*,*N*-dimethlformamide
DMTr	dimethoxytrityl
DNA	2'-deoxyribonucleic acid
dNTP	2'-deoxynucleoside 5'-triphosphate
dU	2'-deoxyuridine
F	2,4-difluorotolene-2'-deoxyrinucleoside
$F^6$2P	6-fluoro-5-methylpyridin-2(*1H*)-one
FAPy	2-amino-6-fluoropyridine
G	2'-deoxyguanidine
HIV	human immunodeficiency virus
HMDS	1,1,1,3,3,3-hexamethyldisilazane
HIV-RT	human immunodeficiency virus reverse transcriptase
HPLC	high performance liquid chromatography
KF	klenow fragment
KF-	exonuclease-free klenow fragment
Klentaq1	large fragment of *Thermus aquaticus* DNA polymerase I
m	multiplet
$m^6$2P	6-methylpyridin-2(*1H*)-one
m^6APy	2-amino-6-methylpyridine
MeCN	actonitrile
MMLV-RT	Moloney murine leukemia virus reverse transcriptase

mtDNA	mitochondrial DNA
NMR	nuclear magnetic resonance spectroscopy
PCR	polymerase chain reaction
Pd	palladium
Ph	phenyl
Phe	phenylalanine
*p*NPE	*para*-nitrophenylethyl
Pol α	calf thymus DNA polymerase α
Pol β	human DNA polymerase β
Pol γ	DNA polymerase γ
pTSA	*para*-toluene sulfonic acid
Q	deoxyadenine analogue without Watson-Crick hydrogen bonding ability
RNA	ribonucleic acid
RT	reverse transcriptase
s	singlet
T	thymidine
t	triplet
T7-	exonuclease-inactive T7 polymerase
TAM	thymidine analogue mutation
Taq	*Thermus aquaticus* DNA polymerase
TBDPS	*tert*-butyldiphenylsilyl
TEA	triethylamine
TEAB	triethylammonium bicarbonate
THF	tetrahydrofuran
TLC	thin-layer chromatography
UV	ultraviolet spectroscopy
Z	deoxyadenine analogue without Watson-Crick hydrogen bonding ability and minor groove N3

ABSTRACT

Synthesis of Pyrimidine *C*-Nucleoside Analogues and Triphosphate Derivatives

Heng Ming Chan

Five pyrimidine *C*-nucleosides were prepared via Heck-type coupling reactions. These derivatives are designed to mimic dC and dU (or T). The minor groove O2 carbonyl in each derivative is replaced by a hydrogen, a fluorine, or a methyl group. The hydrogen-substituted dC analogue was converted into a 2',3'-dideoxynucleoside, which was converted into a 5'-triphosphate derivative. The other two dC analogues were transformed into 5'-triphosphate derivatives immediately after Heck coupling reactions. These analogues will allow an examination of the nature and role of minor groove interactions between incoming triphosphates and various polymerases.

Introduction

Most organisms rely on accurate replication of the individual genome to survive and continue life. It was only 55 years ago when 2'-deoxyribonucleic acid (DNA) was proved to contain the genetic information[1] and that DNA polymerases were responsible for making copies of the genome.[2] DNA replication allows subsequent events such as transcription and translation for cellular construction and biological activities. For over a half century, a continuous flow of research papers have uncovered many details of this enzyme-mediated process, attempting to put all the pieces of the puzzle together. Structural understanding of the DNA base pairs, the duplex and the polymerases provide a basis for our investigation of this mechanism.

Natural DNA base pairs are held together by hydrogen bonds. The heterocycles of natural nucleosides provide specific arrangement of hydrogen bond donors and acceptors, which allows pairing specificity of base pairs. Adenine matches only with thymine (uracil in RNA) via two hydrogen bonds, and guanine is only complementary to cytosine by three hydrogen bonds (Figure 1). These bonds are important for holding together two complementary strands of DNA chains as discovered by James Watson and Francis Crick[3] in 1953, hence the name "Watson-Crick hydrogen bonds." Watson and Crick also proposed that the DNA duplex adopts a right-handed helical structure, where the two strands are oriented in an anti-parallel fashion. The double helix is often described in analogy to a gently twisted ladder, in which the sugar – phosphate backbones on the outside are the two rails of a ladder and the bases inside are the rungs as they are co-planar and stacked on top of each another (Figure 2). Two apparent macroscopic features of a DNA (or RNA) duplex are a major groove and a minor groove. Functional groups of

the heterocycles facing these grooves, respectively, play essential roles in binding and recognition by proteins and small molecules.[4]

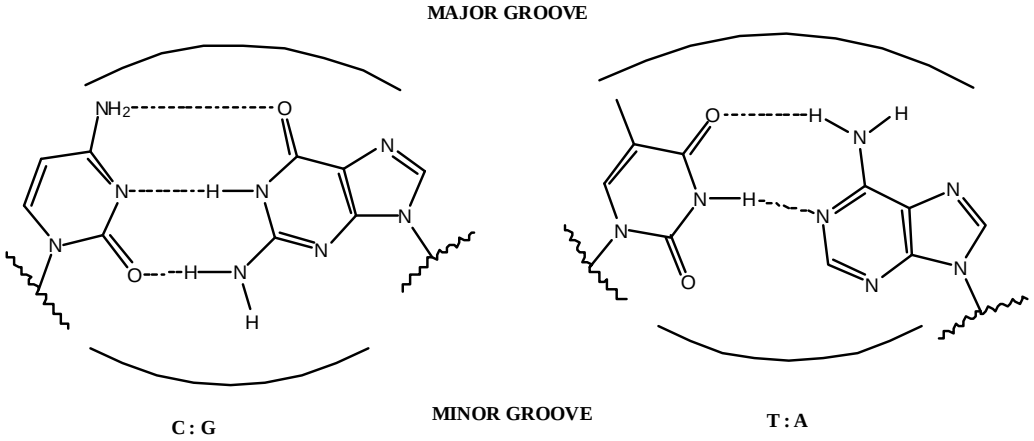


Figure 1. The complementary base pairs of C-G and T-A. The arches above the base pairs encompass the functional groups in the major groove and those below encompass the functional grooves in the minor groove.

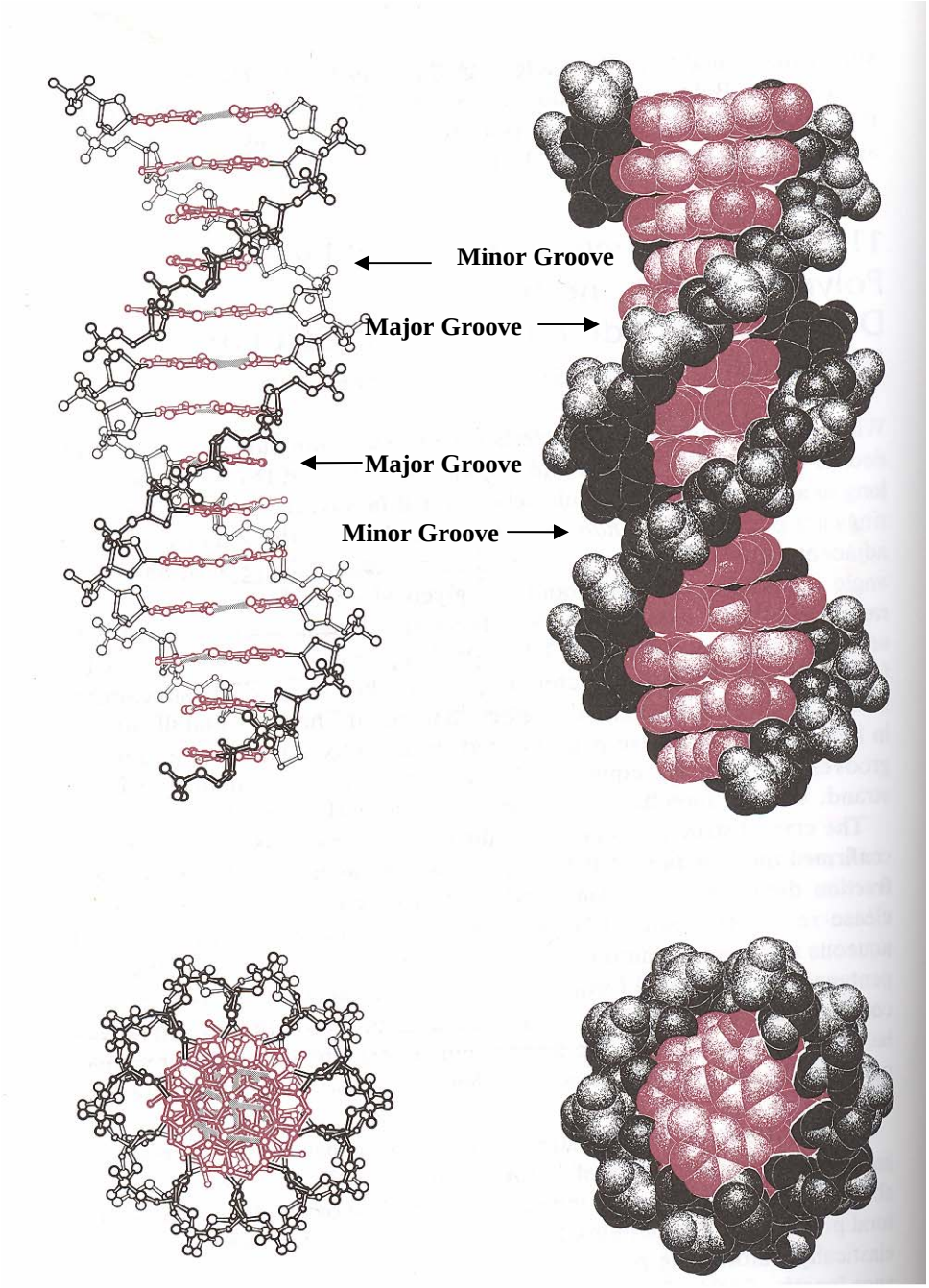


Figure 2. DNA duplex represented in ball-and-stick and space filling representations. Both the side and top views are shown. Minor and major grooves are indicated by arrows. Image adapted from Saenger.[4]

Several crystal structure analyses of DNA polymerases have revealed that many DNA polymerases share the same general multidomain architecture.[5] Polymerases like

Escherichia coli DNA Pol I, T7 DNA polymerase, *Thermus aquaticus* DNA polymerase (*Taq*), and retroviral polymerases such as human immunodeficiency virus reverse transcriptase (HIV-RT) have subdomains described as the palm, fingers and thumb (Figure 3). These studies also suggested that the polymerases undergo structural changes during polymerization. The polymerization process takes place in steps.[6] Initially, the unliganded polymerase has the thumb subdomain in 'closed' position (indicated as E in Figure 4).[7] It moves the thumb to accommodate a nucleic acid duplex to form a binary complex (step 1 in Figure 4).[8] The primer terminus of the DNA duplex is positioned at the priming site (P site) of the active site, allowing the incoming triphosphate to bind at the nucleotide – binding site (N site).[6] Upon binding of an incoming dNTP (step 2), the fingers subdomain change from open conformation to closed by pressing onto the triphosphate (step 3). This conformational change is likely rate determining[9] and seems to conform to an "induced fit model," in which the *correct* nucleotide triggers the enzyme to adopt an active configuration.[10] The polymerase arranges key residues in the active site for stabilizing interactions with the DNA duplex and with the substrate. Many of these contacts include those to the minor groove functional groups. In contrast, an incorrect substrate results in a poor "fit" into the active site and the conformational change is slowed down. A rapid chemistry step follows the conformational change to form a new phosphodiester bond accompanied by the release of pyrophosphate as a byproduct and possibly opening of the fingers subdomain (step 4). In processive polymerization, the polymerase moves onto the next nucleotide on the template by translocating the primer terminus to the P site and continue the catalytic cycle (step 5). If polymerization is distributive, the enzyme dissociates from the elongated primer/template

duplex and the polymerase restarts from binding to a new DNA duplex. The efficiency of the primer terminus translocation determines whether the enzyme will continue polymerizing or fall off from the nucleic acid.[6] It is also associated with the resistance of some mutated reverse transcriptases to their inhibitors (will be discussed later).

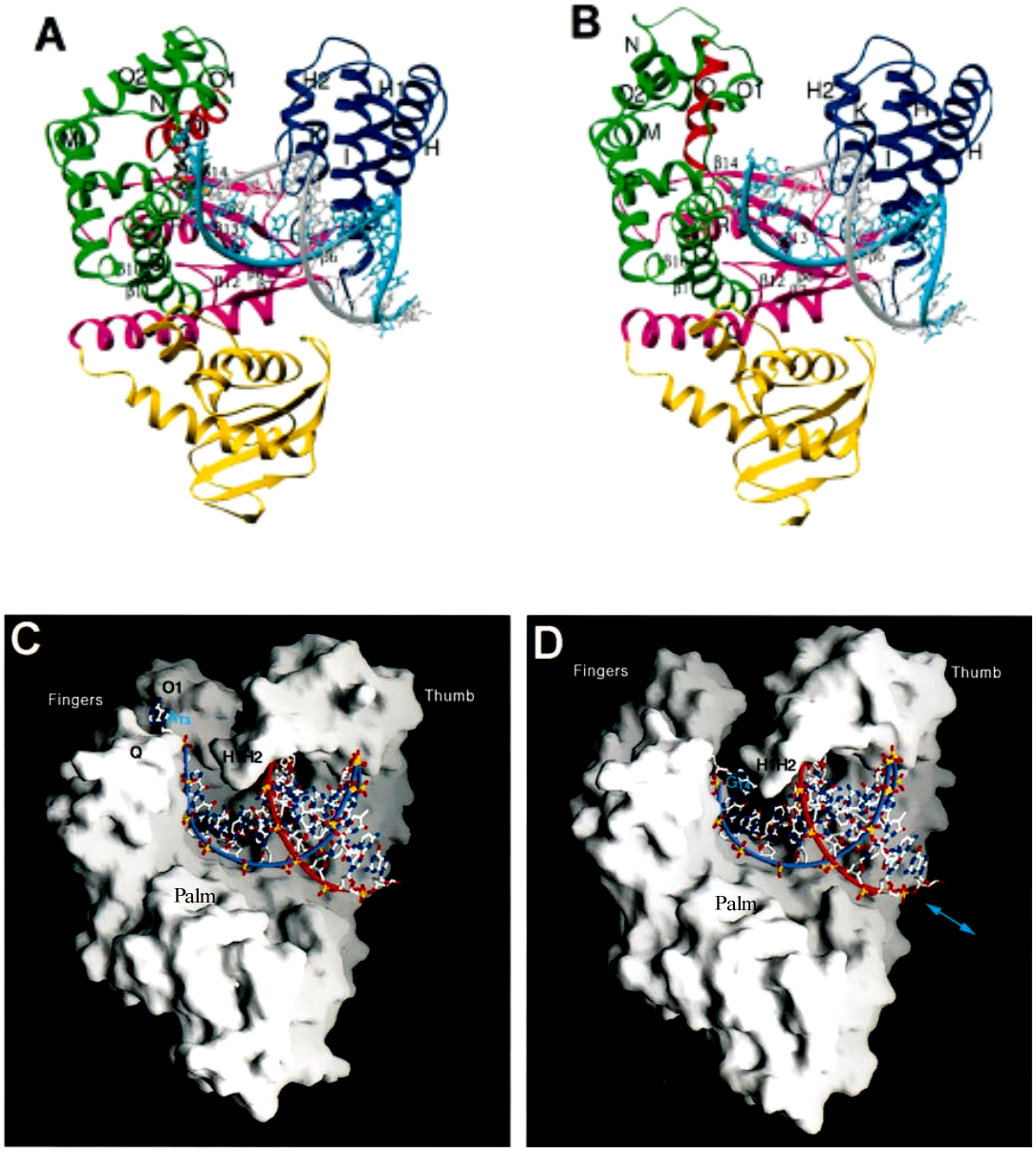


Figure 3. Closed and open forms of the large fragment of *Thermus aquaticus* DNA polymerase I (Klentaq1). A and C: a 'closed' ternary complex structure of Klentaq1 with a DNA duplex and a triphosphate substrate. The fingers domain closes the crevice formed by thumb, palm, and fingers. B and D: a binary complex without the triphosphate

substrate. The fingers domain is in the open form, in which the crevice is clearly visible. Adapted from Li, Korolev and Waksman.[5b]

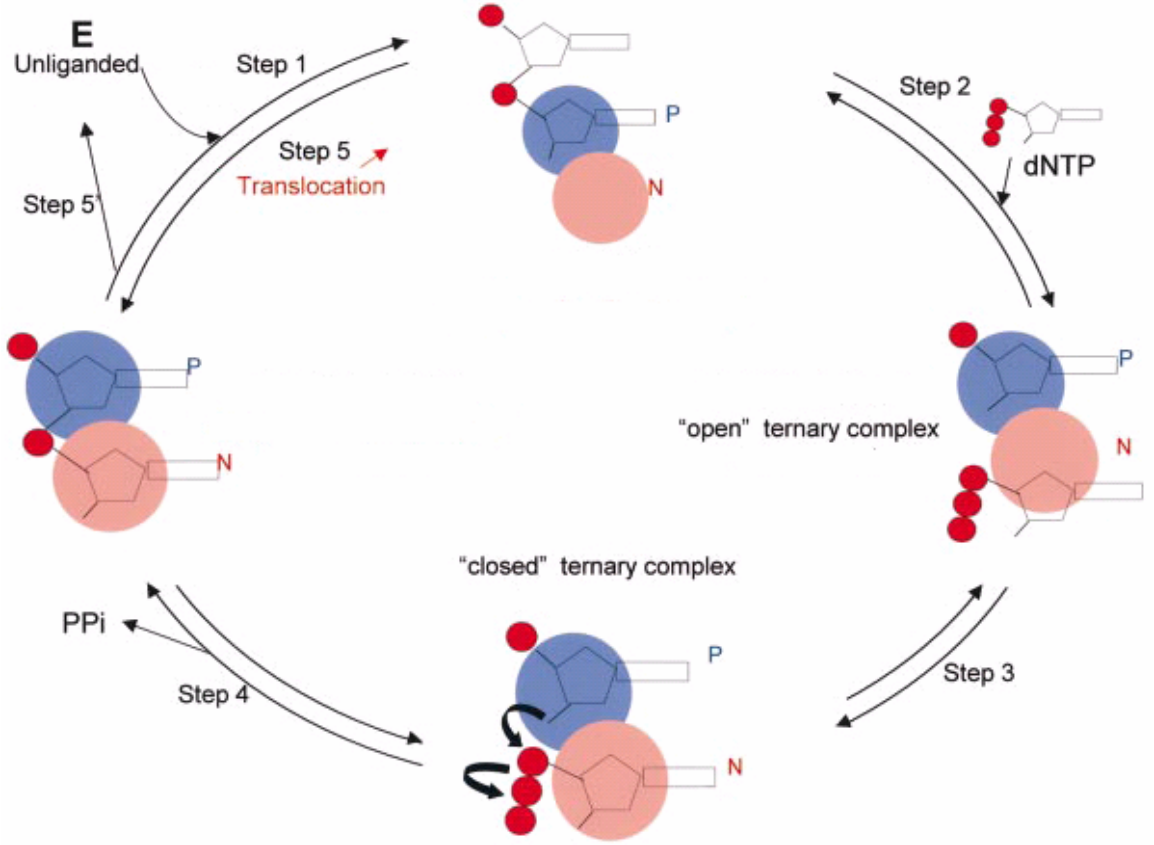


Figure 4. Schematic representation of the stepwise replication process by a polymerase. Only the primer terminus is shown occupying the priming site (P) or the nucleotide-binding site (N). Step 1: the unliganded polymerase (E) binds to a DNA duplex and positions the primer terminus in the P site. Step 2: a nucleotide triphosphate binds in the N site to form an "open" ternary complex. Step 3: the enzyme undergoes conformational changes to form a "closed" ternary complex. Step 4: a chemical step takes place to form a new phosphodiester bond and release inorganic phosphate. Step 5: in processive synthesis, translocation happens to place the new primer terminus in the P site to continue polymerization. Step 5': in distributive synthesis, the polymerase dissociates from DNA to return to the unliganded state. Adapted from Sarafianos et. al.[6]

While most DNA polymerases produce DNA copies of the genome according to a DNA blueprint, some retroviruses such as human immunodeficiency virus (HIV) contain an RNA genome but replicate through a double-stranded DNA intermediate. The enzyme

responsible for such reversed transcription is known as a reverse transcriptase (RT). In the life cycle of a typical HIV, the retrovirus infects host cells, primarily T-helper lymphocytes and macrophages in the immune system. The viral RNA is uncoated in the host's cytoplasm. Reverse transcriptase synthesizes a complementary strand of DNA (DNA transcript) based on the RNA template. Then, it digests the RNA genome and synthesizes a second strand of DNA that is complementary to the first.[11] The double-helical viral DNA (cDNA) is now available for integration into the host genome. Using the cellular resources of the host, new viral RNA and proteins are produced for viral assembly and eventual release of the virion. HIV infection causes acquired immune deficiency syndrome (AIDS) and infected individuals could die from collapsed immune system. Reverse transcription is the most therapeutically targeted stage in the HIV lifecycle because inhibition will prevent the synthesis of cDNA and the ultimate insult.

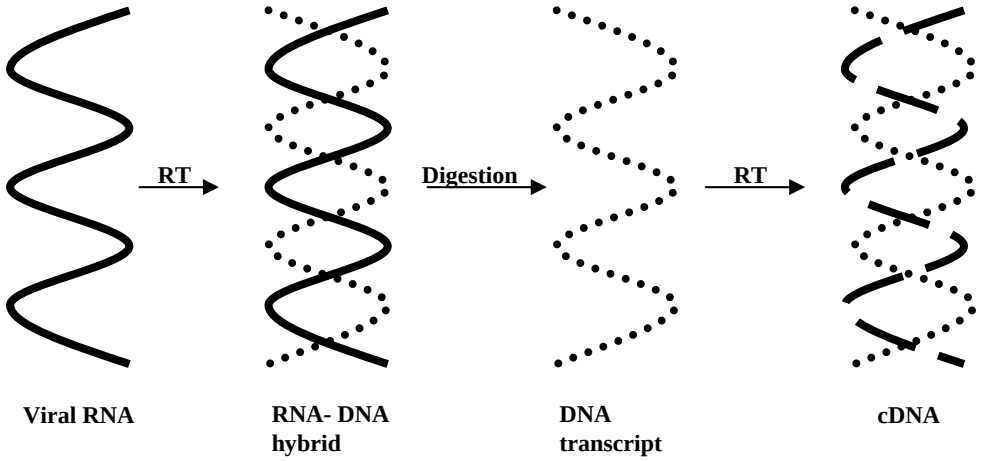


Figure 5. Synthesis of cDNA by RT using an RNA template. A complementary DNA transcript is first synthesized. The template is digested then a second complementary DNA strand is synthesized to make a double stranded DNA copy.

A class of modified nucleosides specifically targets the RT; they are called reverse transcriptase nucleoside inhibitors (RTNIs).[12] Some of them are routinely used in the clinical treatment of HIV-infected individuals. They are 2'-deoxyribonucleoside analogues without the 3'-OH on the ribose moiety. They include zidovudine (AZT),[13] zalcitabine (ddC),[14] didanosine (ddI), stavudine (d4T),[15] lamivudine (3TC), and abacavir (ABC) (Figure 6). Intracellular conversion to 5'-triphosphate derivatives allows them to compete with natural nucleotides for the recognition by reverse transcriptase as substrates. When incorporated into nascent viral DNA chain, the NRTIs fail to continue the elongating chain due to the lack of a 3'-nucleophile, so they are also known as "chain terminators."

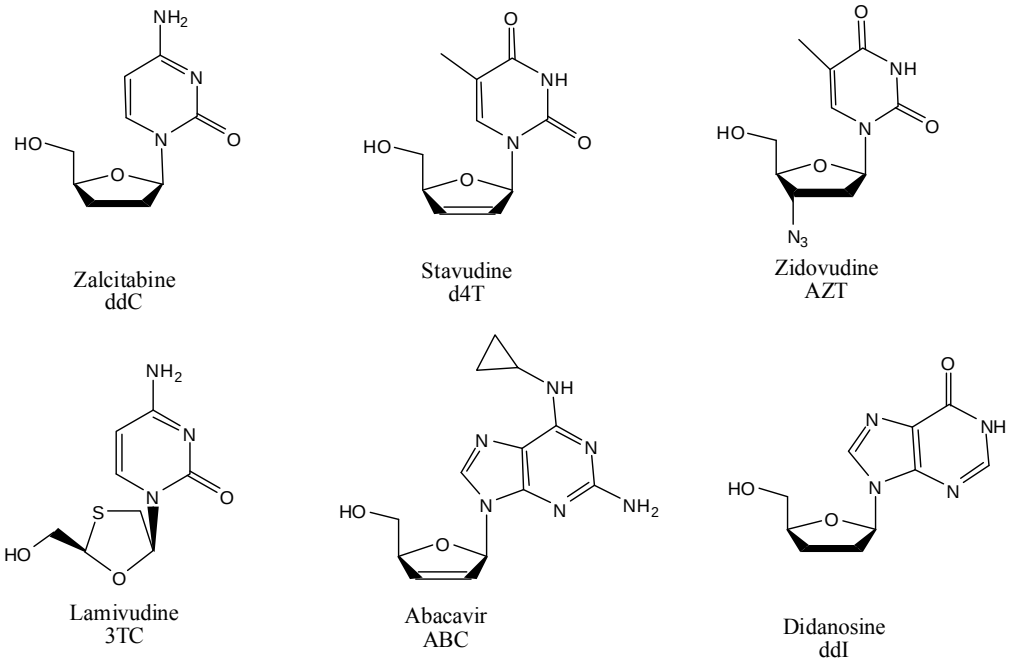


Figure 6. Anti-HIV RT nucleoside inhibitors used in clinical treatment of AIDS.

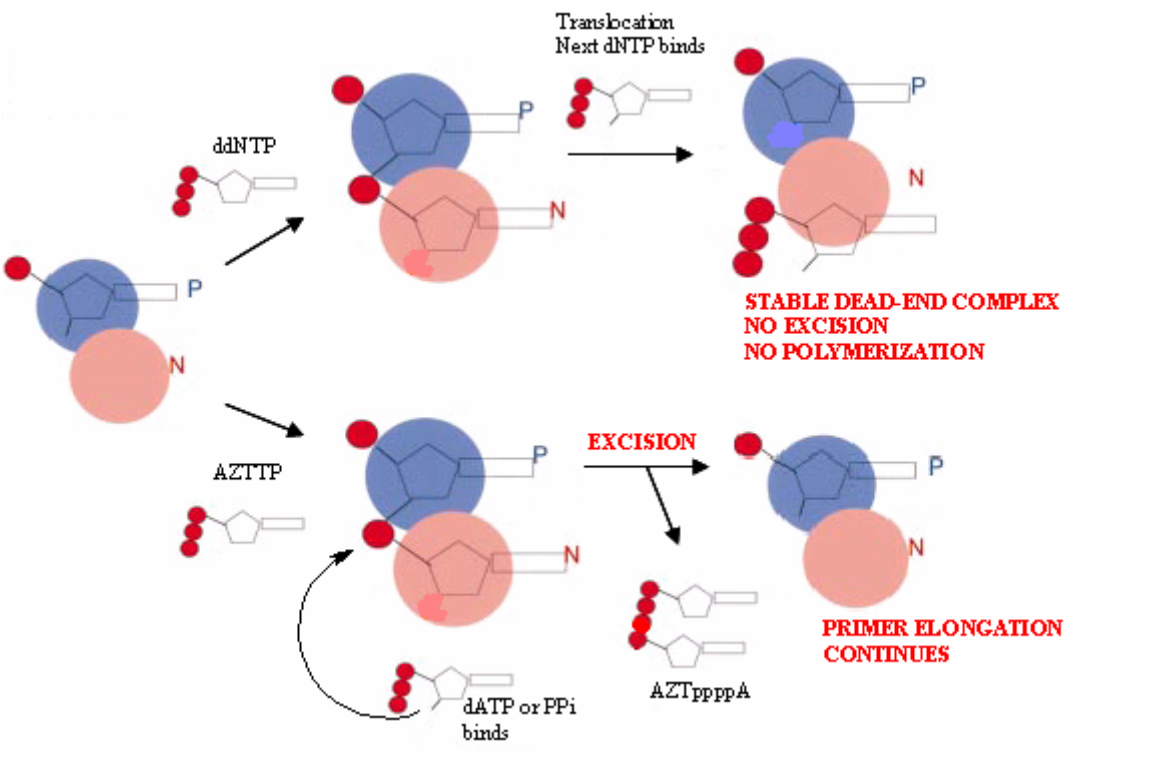


Figure 7. Mechanism of nucleoside chain terminators and TAM-associated excision. Top panel: dideoxynucleotide is incorporated into primer to form a stable dead-end complex. Bottom panel: AZT is excised by dATP or inorganic phosphate and the terminated primer is unblocked. Adapted from Goldschmidt and Marquet.[16]

AZT and other dideoxynucleoside inhibitors take effect by forming a stable dead-end complex with reverse transcriptase (Figure 7 top).[17] Just as in the insertion mechanism discussed earlier, RT inserts the RTNI (ddNTP) into the primer in the N site of the active site. The terminal RTNI phosphodiester translocates to the P site to accommodate an incoming dNTP. The dNTP occupies the N site and the formation of a new phosphodiester bond is precluded by the nature of the RTNI. As a result, the polymerase – DNA complex is trapped in a "dead-end" complex and polymerization is disrupted.

The emergence of resistant strains of HIV, however, has enabled two major phenotypic mechanisms to hamper the efficacy of the drug.[7a,16] The first one is RTNI

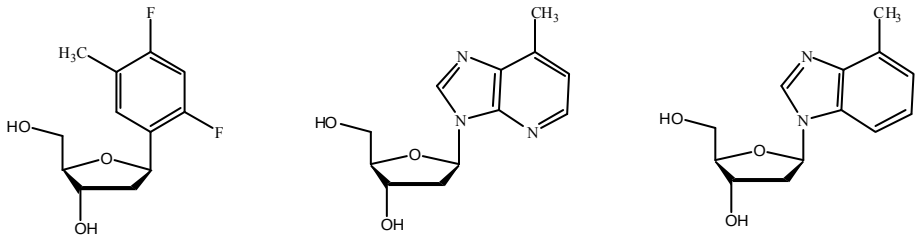
discrimination, in which the mutant reverse transcriptase preferentially incorporates natural nucleotides over synthetic ones. The second involves pyrophosphate or ATP-dependent phosphorolytic excision of the *incorporated* RTNI from the 3'-end of the primer strand. It is primarily associated with AZT, caused by a set of mutations commonly known as thymidine analogue mutations (TAMs) (Figure 7 bottom).[18] In the case of TAM bearing polymerases, AZT is incorporated into the primer in the N site. Because of enhanced affinity for ATP and inorganic pyrophosphate,[16] the mutated RT binds to either one and the pyrophosphate donor attacks the AZT phosphodiester. The AZT phosphodiester is bonded to the ATP to form a dinucleotide polyphosphate or to the pyrophosphate to generate AZT triphosphate. In either case, AZT is excised from the primer terminus and the initially terminated primer could resume elongation. In this regard, patients with resistant strains of HIV show a reduced sensitivity to these thymidine analogues.

Besides the problem of resistance, many existing anti-HIV dideoxynucleosides including AZT, ddC, and ddI have shown cytotoxicity effects. Delayed toxicity most commonly takes the form of peripheral neuropathy, myopathy, or pancreatitis.[19] Kuchta and coworkers[20] discovered a pathway through which AZT monophosphate accumulates and potently inhibits protein and lipid glycosylation in the Golgi complex and endoplasmic reticulum. This results in glycolipid composition change in the cellular membrane and thereby alters the properties of many receptors and enzymes.[21] Anemias that result from killing of blood progenitor cells also possibly involve inhibition of protein glycosylation.[22] Medina and coworkers[23] compared three anti-HIV dideoxy-nucleosides ddC, d4T, and ddI and found significant reduction in cell viability and

mitochondrial DNA (mtDNA) content as well as mitochondrial morphology distortion in treated cultured cells. The damage was proportional to the concentration of the nucleosides. The cytotoxicity effects of three were in the order of ddC > D4T > ddI. It seems obvious from these studies that interference with cellular polymerases or activities adds another challenge to creating clinically practical anti-HIV agents. Scientists have developed a wide range of synthetic nucleosides and studied their interactions with various polymerases and reverse transcriptases, in order to identify a crucial factor to elicit enzyme discrimination.

Kool and coworkers have introduced a class of nonpolar nucleoside analogues and used them to probe minor groove interactions with various polymerases and reverse transcriptases. These analogues mimic as closely as possible the size and shape of natural nucleosides but have little or no Watson-Crick hydrogen bonding capacity (Figure 8). They have been placed in the template of an incipient base pair and used as triphosphate substrates with various polymerases and reverse transcriptases.[24] These enzymes were grouped into three based on the number of minor groove interactions. The first group efficiently replicated base pairs containing F and Z: the exonuclease-free Klenow fragment (KF-), exonuclease-inactive T7 polymerase (T7-), *Taq* polymerase, and HIV-RT. They do not seem to require any contacts between amino acid side chains at the active site and the minor groove side of the incipient base pair. Available crystal structures of some of these enzymes complexed with double stranded DNA and a nucleotide supported this.[25] The second group of polymerases consisted of calf thymus DNA polymerase α (Pol α) and avian myeloblastosis virus reverse transcriptase (AMV-RT), which required a minor groove interaction at either the template base or the

incoming triphosphate. The last group formed by human DNA polymerase (Pol β) and Moloney murine leukemia virus reverse transcriptase (MMLV-RT) failed to replicate any of the nonpolar base pairs, suggesting that these enzymes possibly require minor groove interactions at both the template base and the incoming nucleotide simultaneously for efficient replication. This study of nucleotide-protein interactions at the insertion step demonstrated the plausibility of using base-modified analogues to target HIV-RT but leave out eukaryotic polymerases such as Pol α and Pol β.



F **Q** **Z**

Figure 8. Kool's hydrophobic isosteres.

More importantly, these nonpolar isosteres have established that shape complementarity (avoidance of steric clashes) plays a critical role in DNA replication.[26] Kool and coworkers demonstrated that the hydrophobic analogue F is inserted across from A with an efficiency nearly the same as that of a T-A pair.[27] F is also replicated efficiently and selectively against two adenine shape mimics Z and Q *in vitro* and *in vivo*, respectively.[25] Kool advocates that shape complementarity is a stricter requirement than the need for Waston-Crick hydrogen bonds in polymerase-catalyzed base pair synthesis.[28] He reinforces this idea using a steric exclusion hypothesis to explain how a

geometrically disfavored base pair would result in exclusion from the active site of a polymerase (Figure 9). The active site of the polymerase is defined by the enzyme itself and the incipient base pair enclosed. With natural nucleotides alone, there may be four possible sizes and shapes for the active site cavities during each successive nucleotide insertion, one per each of the four natural bases. The polymerase excludes purine – purine pairs simply because they are too large to fit in the binding pocket together (case 2). The enzyme also avoids inserting an incorrect pyrimidine opposite a given purine because of a proton – proton steric clash when C is across from A or T from G (case 3).

In the case of a pyrimidine – pyrimidine mismatch, both relatively small nucleotides can fit in the binding pocket, but solvation effects prevent such misinsertion (case 4). Both pyrimidines can form hydrogen bonds with water molecules through their Watson-Crick hydrogen bonding groups. The solvating waters can exchange with each other, but they will effectively never be lost in the aqueous environment. They are energetically favorable with a calculated free energy range of -3.7 to -14.5 kcal/mol.[29] A single water adds about 3 to 4 Å of steric bulk to the Watson – Crick edge of each pyrimidine.[30] This makes the effective size of a pyrimidine large enough to be excluded from the active site. The pyrimidines will not exchange the waters to bind with each other because of two possible reasons: 1) hydrogen bonds are strong (several kcal/mol) so desolvation of two bases would be highly costly, and 2) even if desolvation did occur, the two bases would have to achieve complementarity by altering the geometry (as in T-T wobble pair), resulting in steric clashes with the walls of the enzyme active site. Similar effects would be operative for pyrimidine – purine and purine - purine mispairs. Thus,

steric effects enforce shape complementarity of a base pair in the active site and solvation effects enhance these steric effects.

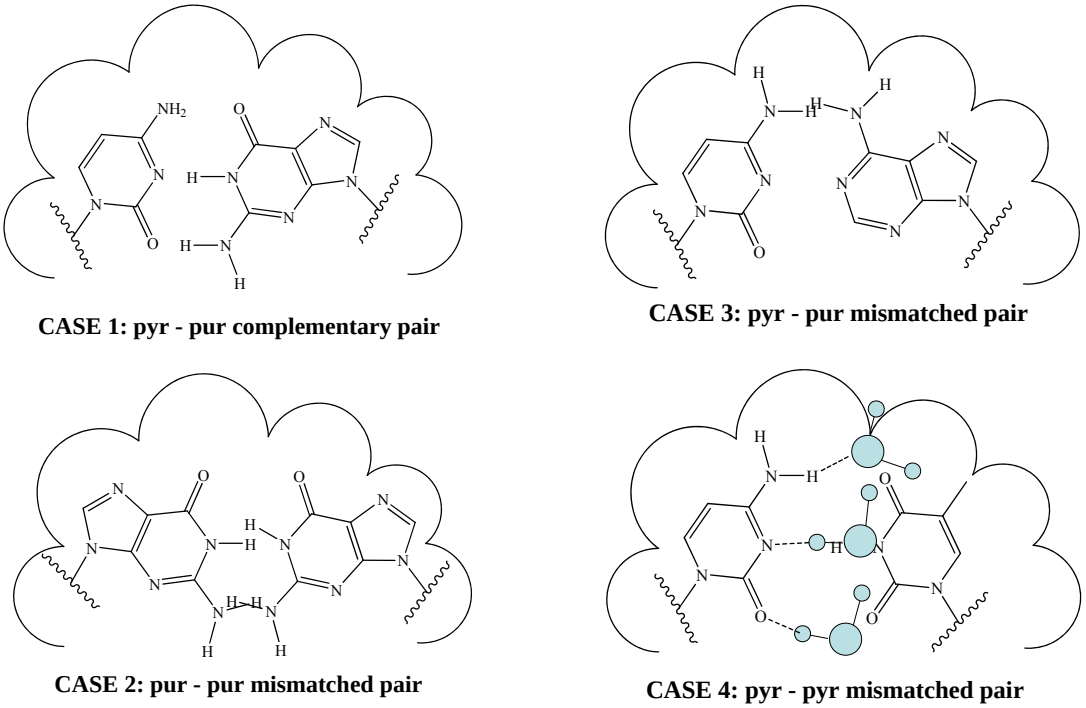
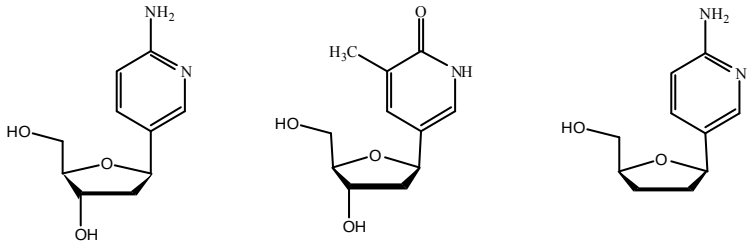


Figure 9. Representation of matched and mismatched pairs in the active site of the polymerase. Steric clashes in cases 2-4 could lead to exclusion of mismatches. Water molecules are shown hydrogen bonded to one pyrimidine in 4.[28]

The McLaughlin group reported two pyrimidine nucleoside analogues, d*C and d*T (Figure 10) that also mimicked the shape of a natural base, but selective removal of the minor groove functional group seemed to cause inhibitory effects toward polymerases KF and *Taq*.[31] Two of the Watson-Crick hydrogen bonds of the pyrimidine bases are reserved to take advantage of the high cost of desolvation in the case of mismatched pairs. Direct elimination of the minor groove O2 carbonyl resulted in attendant changes of the hydrogen-bonding character of the N3-nitrogen. The presence of this analogue in DNA results in significant duplex destabilization.[32] By preparing the corresponding *C*-

nucleosides,[33] however, the correct tautomeric forms of the bases are maintained for proper Watson-Crick-like hydrogen bonds. Syntheses of *C*-nucleosides have been well explored.[34] Most naturally occurring *C*-nucleoside compounds are antibiotics and many also possess antiviral and/or antitumor properties.[35]



d*C **d*T** **dd2APy**

Figure 10. Nucleoside analogues used in McLaughlin's studies.

The triphosphate derivatives d*CTP and d*TTP were found to inhibit the *Taq* polymerase under standard PCR conditions.[31] The experiments were repeated with KF and progressive concentration-dependent suppression of full-length products was confirmed. These results indicated that pyrimidine nucleotides lacking the 2-keto group are poor substrates for *Taq* and KF. The keto group seems to assist interactions between the incoming triphosphate and some amino acid side chains in the active site of the polymerase in order to correctly position the incoming triphosphate for proper alignment with the primer terminus. Without the keto group, however, the unnatural triphosphate may be placed in the active site in a position where necessary conformational change of the enzyme fails to permit formation of a phosphodiester bond and release of

pyrophosphate. Therefore, the deletion of the minor groove carbonyl group seems to have given rise to the inhibitory effect of these derivatives.

Upon this finding, the McLaughlin group went on the produce a dideoxynucleoside triphosphate dd2APyTP (Figure 10) that showed substrate activity with HIV RT but discriminated against polymerases α, β, and to some extent γ.[36] The design of the cytidine analogue d2APy combines two advantageous features: the 2',3'-dideoxy carbohydrate moiety will ensure chain termination and the heterocyclic nucleobase without the minor groove O2 carbonyl group may elicit enzyme discrimination. In a 45-minute time course assay, no significant elongation of the primer strand was observed for DNA Pol α or β at dd2APyTP concentration as high as 1mM, while it only took low μM concentration of the natural nucleotide for the same elongation to occur. On the other hand, mitochondrial polymerase (Pol γ) extended roughly 11% of the primer in the presence of 100 μM dd2APyTP and HIV RT 50% under the same conditions. Since most cytotoxicity instances involve the mitochondrial polymerase using the chain terminator for mitochondrial DNA replication, the enzyme was incubated with dd2APy-incorporated DNA to estimate the cytotoxicity effect. The rate of Pol γ removing the dideoxynucleotide by exonuclease activity was commensurate with that of 3TC and significantly faster than that for ddC.[37] That means the dideoxynucleotide cytotoxicity level was well tolerated if not better than the clinically approved chain terminators. Thus, dd2APy presents a selective antiviral prodrug with limited cytotoxicity. Kinetic data remain yet to be obtained to determine the insertion efficiency for this particular substrate.

Since the extent of shape alteration arising from the complete deletion of the pyrimidine minor groove functional group remains unclear, substitutions in the minor

groove region with fluorine and a methyl group may help examine the nature of minor groove interactions between the HIV-RT and the nucleotide. We have replaced the O2 carbonyl with a fluorine (FAPy and $F^6$2P) or a methyl group (m^6APy and $m^6$2P), in addition to the previously mentioned hydrogen-substituted analogues d2APy and d*T (Figure 11). The methyl group in the fifth position of thymine will be deleted for the simplicity of synthesis because its distance from the minor groove region should diminish any effect on minor groove interaction. The previously discussed thymidine analogue, difluorotoluene (F),[23,25,26] also contains a fluorine substitution in the minor groove region. Its specific insertion across adenine by KF might have been made possible by fluorine-hydrogen bonds, besides the fact that F is a good shape mimic of T. In addition, fluorine may also replicate the electron-withdrawing effect of a carbonyl group, making it an electronic mimic of the carbonyl. It remains to be determined whether the F-A pair is a result of shape complementarity, electronic effect, fluorine-hydrogen bonding, or a combination of all. Nonetheless, F represents a successful example of polymerase substrates FAPy and $F^6$2P nucleosides may possess the same advantages. Furthermore, FAPy nucleotide has been incorporated into a DNA duplex to probe interactions with a minor groove binding molecule.[38] The presence of single fluorine in a DNA duplex appears to promote helix stabilization and native-like stabilization occurs when both fluorines are present. These interactions are likely the sum of fluorine-hydrogen bonding and dipole-dipole or dipole-quadrupole interaction. The methyl substitution in the minor groove may be compared to a thiocarbonyl replacement in the same position[39] because they both increase the size of the thymine base by a small increment (the thiocarbonyl is 0.45 Å longer than the carbonyl). The 2-thiocarbonyl-substituted thymidine analogue

demonstrated that the modification stabilizes a duplex better than the natural congener and that it base pairs with A with high selectivity. The methyl group of m^6APy and $m^6$2P nucleosides might also provide similar base pairing selectivity by size exclusion and possibly hydrophobic interactions with the polymerase active site. Even if the methyl-substituted nucleotide is not a polymerase substrate, the results may have implications on the size or electronic requirement for base pair synthesis.

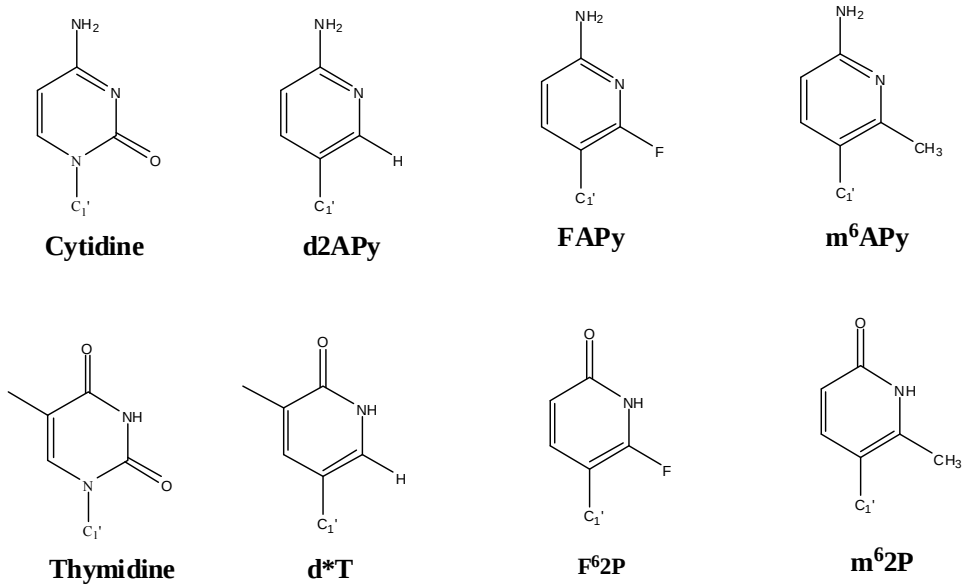


Figure 11. Modified nucleobases of target molecules compared to natural pyrimidine bases.

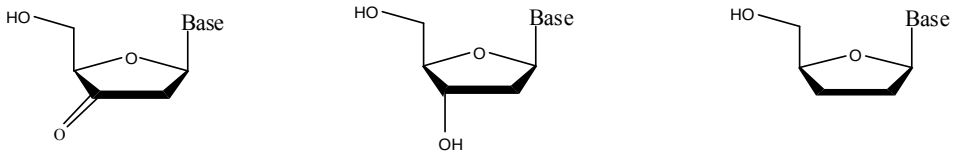


Figure 12. Product of Heck coupling reactions differ from natural and dideoxy nucleosides in the substitution at the 3' position of the carbohydrate.

The current work will present a convergent synthetic scheme for the precursors of the dideoxynucleosides. The precursors are the products of a popular *C*-nucleoside

preparative approach – the Heck coupling reaction (Figure 12). Since they also lack the 3'-OH, they are technically chain terminators just as the dideoxy derivatives. The presence of a carbonyl at the 3' position of the carbohydrate moiety could be constructive because it may better imitate the 3'-OH group than a hydrogen in terms of bond length and polarity, although the geometry of the 3'-carbon is sp^2 in place of sp^3 (Figure 12). The 3'-OH group was found important during base pair synthesis by KF - a Phe762 side chain of the active site makes a snug fit that constrains the deoxyribose into the correct position for the 3'-OH to make contact.[40] This interaction contributes to the setup of the three-dimensional arrangement for proper conformational change of the enzyme just prior to phosphodiester bond formation. The triphosphate syntheses of dd2APy and Heck coupling products with bases of FAPy and m^6APy will be included in the end of this paper. These triphosphate derivatives will allow future research on their polymerase activities and chain-terminating ability.

References

1 Avery, O., MacLeod, C., & McCarty, M. *Journal of Experimental Medicine* **1944**, 79, 137-157.

2 Kornberg, A. & Baker, T. A. *DNA Replication*, 2nd ed., W. H. Freeman: New York, **1992**.

3 (a) Watson, J. D.; Crick, F. H. *Proc. Roy. Soc. (London) Ser. A* **1954**, 223, 80-93. (b) Watson, J. D.; Crick, F. H. C. *Nature* **1953**, 171, 737-738.

4 Saenger, W. *Principles of Nucleic Acid Structure*; Springer-Verlag: New York, **1984**.

5 (a) Li, Y., Kong, Y., Korolev, S. & Waksman, G. *Protein Science* **1998**, 7, 1116-1123. (b) Li, Y. Korolev, S. & Waksman G. *EMBO Journal* **1998**, 17, 24, 7514-7525. (c) Brautigam, C. A. & Steitz, T. A. *Current Opinion in Structural Biology* **1998**, 8,54-63. (d) Brieba, L. G., Kokoska, R. J., Bebenek, K., Kunkel, T. A. & Ellenberger, T. *Structure* **2005** 13, 1653-1659. (e) Ding, J., Das, K., Hsiou, Y., Sarafianos, S. G., Clark Jr., A. D., Jacobo-Molina, A., Tantillo, C., Hughes, S. H. & Arnold, E. *J. Mol. Biol.* **1998**, 284, 1095-1111.

6 Sarafianos, S. G., Clark, Jr. A. D., Das, K., Tuske, S., Birktoft, J. J., Ilankumaran, P., Ramesha, A. R., Sayer, J. M., Jerina, D. M., Boyer, P. L., Hughes, S. H. & Arnold, E. *EMBO Journal* **2002**, 21, 23, 6614-6624.

7 (a) Rodgers, D.W., Gamblin, S. J., Harris, B. A., Ray, S., Culp, J. S., Hellmig, B., Woolf, D. J., Debouck, C. & Harrison, S. C. *Proc. Natl Acad. Sci. USA* **1995**, 92, 1222-1226. (b) Hsiou, Y., Ding, J., Das, K., Clark, A. D., Jr, Hughes, S. H. & Arnold, E. *Structure* **1996**, 4, 853-860.

8 (a) Jacobo-Molina,A. et al. *Proc. Natl Acad. Sci. USA* **1993**, 90, 6320-6324. (b) Ding, J., Das, K., Hsiou,Y., Sarafianos, S. G., Clark, J., A. D., Jacobo-Molina, A., Tantillo, C., Hughes, S.H. & Arnold, E. *J. Mol. Biol*. **1998**, 284, 1095-1111.

9 Kuchta, R. D., Benkovic, P. & Benkovic, S. J. *Biochemistry* **1988**, 27, 6716-6725.

10 (a) Patel, S. S., Wong, I. & Johnson, K. A. *Biochemistry* **1991**, 30, 511-525. (b) Wong, I. Patel, S. S. & Johnson, K. A., *Biochemistry* **1991**, 30, 526-537. (c) Washington, M. T., Prakash, L. & Prakash, S. *Cell* **2001**, 107, 7, 917-927.

11 Stryer, L. *Biochemistry*, 4th ed., W. H. Freeman & Co.: New York, **1995**.

12 Goody, R. S., Mu″ ller, B. & Restle, T. *FEBS Lett.* **1991** 291**,** 1–5.

13 Chen, B., Quinlan, S. L., & Reid, J. G. *Tetrahedron Letters* **1995**, 36, 44, 7961-7964.

14 Chen, C. & Cheng, Y. *J Biol. Chem.* **1992**, 267, 2856–2859.

15 Huang, P., Farquhar, D. & Plunkett, W. *Journal of biological Chemistry* **1992**, 267, 4, 2817-2822.

16 Goldschmidt, V. & Marquet, R. *Int. J. Biochem. Cell Biol.* **2004**, 36, 1687-1705.

17 Sluis-Cremer, N., Arion, D., Parikh, U., Koontz, D., Schinazi, R. F., Mellors, J. W. & Parniak, M. A. *J. Biol. Chem.* **2005**, 280, 32, 29047-29052.

18 (a) Arion, D., Kaushik, N., McCormick, S., Borkow, G. & Parniak, M. A. *Biochemistry* **1998**, **37,** 15908–15917. (b) Meyer, P. R., Matsuura, S. E., So, A. G. & Scott, W. A. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, **95,** 13471–13476. (c) Meyer, P. R., Matsuura, S. E., Mian, A. M., So, A. G. & Scott, W. A. *Mol. Cell* **1999**, **4,** 35–43.

19 (a) Faulds, D.; Brogden, R. N. *Drugs* **1992**, *44*, 94-116. (b) Whittington, R.; Brogden, R. N. *Drugs* **1992**, *44*, 656-683. (c) Wilde, M. I.; Langtry, H. D. *Drugs* **1993**, *46*, 515-578. (d) Adkins, J. C.; Peters, D. H.; Faulds, D. *Drugs* **1997**, *53*, 1054-1080.

20 Hall, E. T., Yan, J., Melancon, P. & Kuchta, R. D. *J. Biol. Chem.* **1994**, 269, 20, 14355-14358.

21 (a) Bremer, E. G., Hakomori, S.-I., Bowen-Pope, D. F., Raines, E., & Ross, R. *J. Biol. Chem.* **1984**, 259, 6818-6825. (b) Kreutter, D., Kim, J. **Y.** H., Goldenring, J. R., Rasmussen, H., Ukomadu, C., DeLorenzo, R. J. & Yu, R. K. *J. Biol. Chem.* **1987**, 321, 726-738.

22 Sommadossi, J.-P., Carlisle, R., Schinazi, R. F., & Zhou, Z. *Antimicrob. Agents Chemother.* **1988**, 32, 997-1001.

23 Medina, D. J., Tsai, C., Hsiung, G. D. & Cheng, Y. *Antimicrob. Agents Chemother.* **1994**, 1824-1928.

24 Morales, J. C. & Kool, E. T. *JACS* **2000**, 122, 6, 1001-1007.

25 (a) Huang, H., Chopra, R., Verdine, G. L., Harrison, S. C. *Science* **1998**, 282, 1669-1675. (b) Doublie, S., Tablor, S., Long A. M., Richardson, C. C., Ellenberger, T. *Nature* **1998**, 391, 251-258. (c) Eom, S. H., Wang, J., Steitz, T. A. *Nature* **1996**, 382, 278-281.

26 (a) Guckian, K. M., Morales, J. C. & Kool, E. T. *J. Org. Chem.* **1998**, 3, 9652-9656. b) Morales, J. C. & Kool, E. T. *Nature Structural Biology* **1998**, 5, 11, 950-954. c) Delaney, J. C., Henderson, P. T., Helquist, S. A., Morales, J. C., Essigmann, J. M. & Kool, E. T. *PNAS* **2003**, 100, 8, 4469-4473.

27 Morales J. C. & Kool, E. T. *JACS* **1999**, 121, 2323-2324.

28 Kool, E. T. *Biopolymers* **1998**, 48, 1, 3-17.

29 Gao, J. *Biophys. Chem.* **1994**, 51, 253-261.

30 Gao, J. *Biophys. Chem.* **1994**, 51, 253-261.

31 Guo, M., Hildbrand, S., Leumann, C. J., McLaughlin, L. W. & Waring, M. J. *Nucleic Acids Research* **1998**, 26, 8, 1863-1869.

32 Rajur, S. B. & McLaughin, L. W. Unpublished results.

33 (a) Rajur, S. B. & McLaughlin, L. W. *Tetrahedron Lett*. **1992** 33, 6081–6084. (b) Hsieh, H. & McLaughlin, L. W. *J. Org. Chem*. **1995**, 60. 5356–5359.

34 Wu, Q. & Simons, C. *Synthesis* **2004**, 10, 1533-1553.

35 (a) Townsend, L. B. Chemistry of Nucleosides and Nucleotides. Plenum Press: New York, **1994**, 421-535. (b) Miller, P. S. *Antisense/Antigene Oligonucleotides*, In Bioorganic Chemistry-Nucleic Acids. Hecht, S. M., Ed. Oxford University Press: New York & Oxford, **1996**, 347-374. (c) Guntaka, R. V., Varma, B. R. & Weber, K. T. *Int. J. Biochem. Cell Biol.* **2003**, 35, 22. (d) Navarre, J., Guianvarch, D., Giorgio, A. F., Condom, R. & Benhida, R. *Tetrahedron Lett.* **2003**, 44, 2200. (e) Togo, H., He, W., Waki, Y. & Yokoyama, M. *Synlett.* **1998**, 700.

36 Fraley, A. W., Chen, D., Johnson, K. & McLaughlin L. W. *JACS* **2003**, 125, 616-617.

37 Feng, J. Y.; Johnson, A. A.; Johnson, K. A.; Anderson, K. S. *J. Biol. Chem.* **2001**, *276*, 23832-23837.

38 Sun, Z. & McLaughlin, L. W. *JACS* **2007**, 129, 2531-12536.

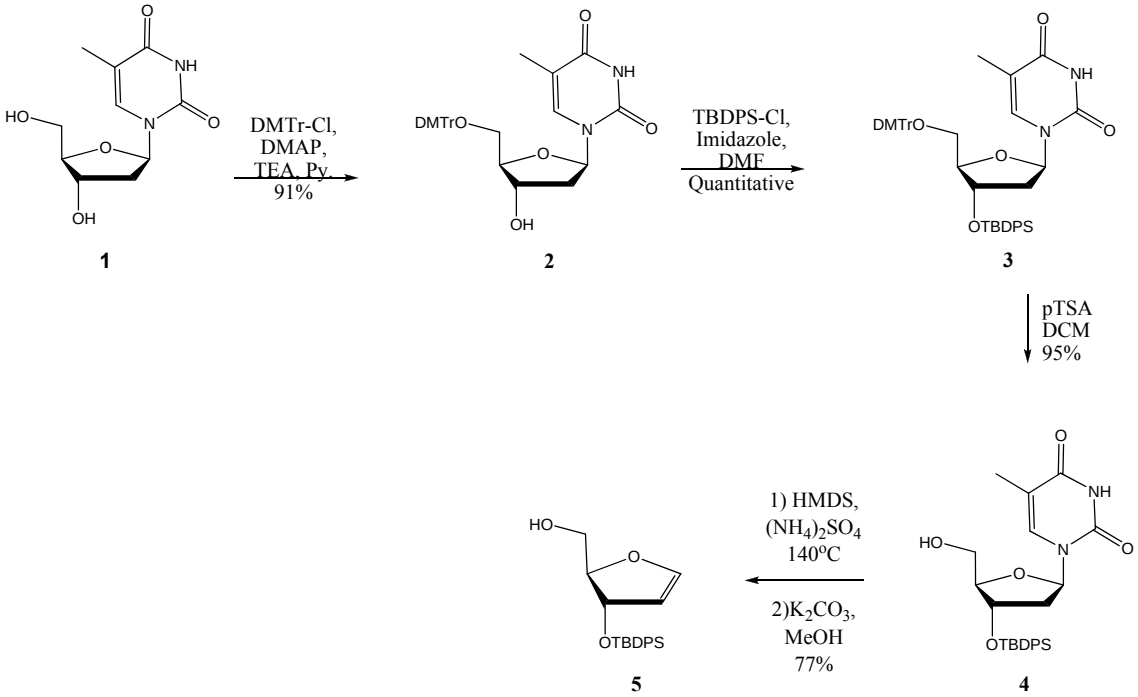
39 Sintim, H. O. & Kool, E. T. *JACS* **2006**, 128, 396-397.

40 Astatke, M., Grindley, N. D. F. & Joyce, C. M. *J. Mol. Biol.* **1998**, 278, 147-165.

Results and Discussion

The synthesis of the *C*-nucleosides involves a Heck-type coupling reaction[1] to connect a sugar bearing a double bond and an iodo-heterocycle. The coupling reaction generates a carbon-carbon glycosidic bond, which helps maintain the correct tautomeric form of the heterocycle for Watson-Crick-like hydrogen bonding. The versatility of this palladium-mediated Heck-type coupling reaction provided convergence of these syntheses and convenience in preparation. The same double-bond-bearing carbohydrate can react with a number of similar iodo-heterocycles under the same conditions to produce a series of *C*-nucleosides.[2] Each reaction yielded only the β anomer in part because a bulky silyl protecting group on the carbohydrate precludes addition of the heterocycle on the bottom face of the double bond. The sugar portion was prepared as follows (Scheme 1).[3]

Scheme 1



Preparation of the Carbohydrate

The goal was to make the α face of carbohydrate sterically hindered. Here we employed extremely bulky and different protecting groups to provide selectivity and stability. Beginning with commercially available thymidine, the stereocenters on the carbohydrate are already set. Taking advantage that a primary hydroxyl group is more reactive than a secondary one, we protected the 5'-hydroxyl group with 1.1 equivalent of DMTr-Cl. After activation with DMAP, the dimethoxytrityl group was installed mostly on the 5'-O position, although a small amount of thymidine simultaneously protected at the 5'- and 3'-O positions was also observed. The excess DMTr reagent was quenched using methanol. For workup, either dichloromethane or diethyl ether could be used in combination with water. When dichloromethane was used, all DMTr-containing compounds remained in the organic layer. When diethyl ether was used, however, all the DMTr-containing compounds were in the organic layer but the desired 5'-O protected thymidine could be precipitated out of the crude mixture. The pure, white solid product (**2)** was filtered. Sometimes sodium sulfate triggered the precipitation during drying after extraction. The drying agent could be removed and washed with dichloromethane to redissolve any trapped product. When precipitation did not happen at all, the mixture was separated on a silica gel column using a small percentage of triethylamine to protect the acid-sensitive dimethoxytrityl.

In the second step, *tert*-butyldiphenylsilane was chosen to protect the 3'-OH for a couple of reasons. First, the protecting group is extremely bulky so the bottom face of the sugar will be blocked completely to allow coupling reaction to occur exclusively on the other face. Second, this silane group is acid stable so the 5'-O-DMTr group can be

removed selectively as it is in the next step. The workup of this reaction was often a simple dichloromethane and water extraction; no column was necessary as long as the starting material was pure and completely consumed.

To remove the 5'-O-DMT group in the third step, we often used *p*-toluenesulfonic acid. Although this reaction should be done within a half hour according to the reported procedure, *p*-toluenesulfonic acid often needed to be added in a few portions over a time period of up to two hours. It was found that adding a small amount of methanol initially helped speed up the reaction because methanol quenches the DMT cations to generate DMTr-OCH_3, favoring deprotection. A couple of aqueous base washes were often performed to neutralize any acid before column purification; the silane group is unstable in acid over extended period of time. For the column, ethyl acetate and hexanes provided a better solvent system than methanol and dichloromethane.

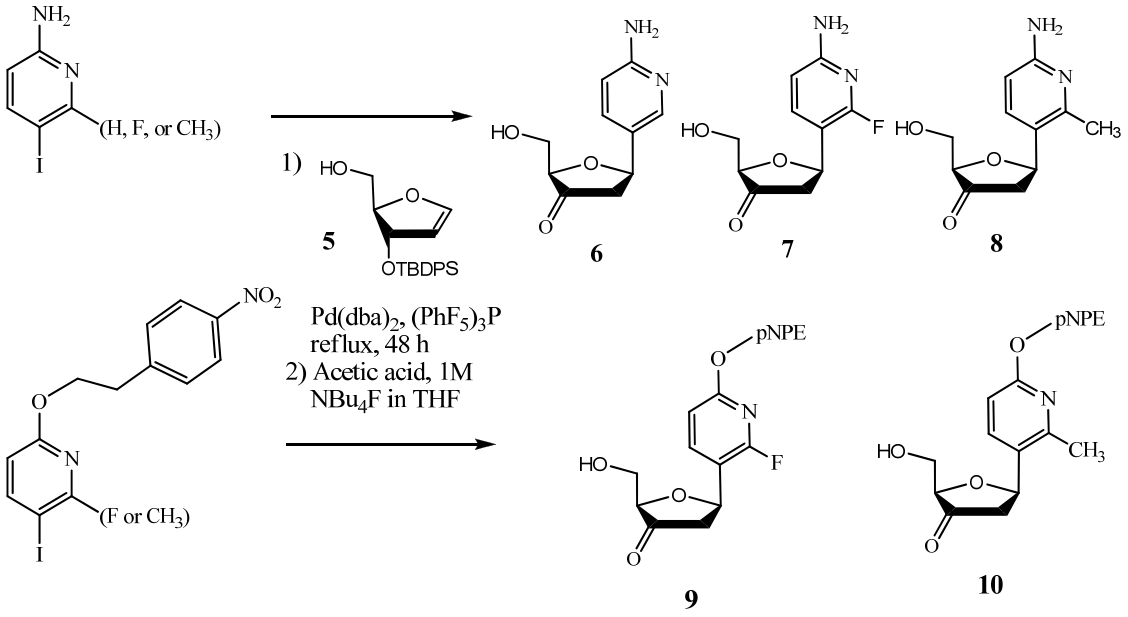
For the last step to synthesize the sterically hindered sugar **5**, thymine was eliminated as a fully silylated heterocycle. 1,1,1,3,3,3-hexamethyldisilazane (HMDS) in the presence of ammonium sulfate reacted with all the oxygen functional groups on thymine as oxygen has strong affinity for silicon. Such protection resulted in a positively charged glycosidic nitrogen. Removal of a 2'-proton cleaves the glycosidic bond and relieves the positive charge of that nitrogen. This elimination reaction was very temperature-sensitive. It required a reaction temperature of at least 140°C. When the mixture was refluxed at only 130°C, not all protected base was removed. The workup for this reaction was modified to improve yield. The old procedure called for a water wash after evaporating the HMDS. This caused serious emulsion and the separation of the two layers was overly time-consuming. This wash was omitted completely. To remove the 5'-

O-silyl group on the sugar moiety, methanol was added directly to the concentrated crude mixture to create a white suspension. The solid was possibly from HMDS precipitate in methanol. Fine powder of K_2CO_3 was added and the reaction was over within half an hour. The old procedure called for a brine wash after removing methanol and it caused the same problem as did the earlier water wash. This wash was omitted as well. The excess K_2CO_3 and the precipitate were filtered together through a thin layer of Celite. The filtrate was a clear solution and concentrated for column purification immediately. The product was easily eluted with straight dichloromethane. The 3'-O-TBDPS group was intact.

Heck-Coupling Reactions

While 2-amino-5-iodopyridine was commercially available, the other iodo-aromatic compounds were synthesized by a former graduate student, Dr. Zhenhua Sun.[3] An undergraduate student, William Issa, reproduced these procedures and contributed the heterocycle derivatives to the following coupling reactions. The carbonyl of uracil analogues was protected as *p*-nitrophenylethylether (*p*NPE).

Scheme 2



The carbohydrate is coupled with various iodo-substituted heterocycles to generate the corresponding C-nucleosides using a palladium catalyst (Scheme 2 & Figure 13). The catalyst mixture is usually freshly prepared by combining palladium (0) and pentafluorotriphenylphosphine in acetonitrile. The color of the mixture turns from black to orange brown, indicating the readiness of the catalyst. Oxidative insertion of the metal into the carbon-iodide bond of the heterocycle initiates the coupling cycle. Palladium is activated from the ground state to a palladium (II) species. Addition of the carbohydrate allows π - complexation between the Pd (II) complex and the olefin. The electron-rich heterocycle tends to bond with the relatively electron poor anomeric carbon and palladium (II) tends to go onto the relatively electron rich 2'-carbon of the carbohydrate. The resulting complex is sometimes called a sigma-adduct as a new sigma bond is formed. The sigma-adduct then decomposes preferably by syn-elimination of a β-hydride and palladium, generating the coupled product with a new π bond between the 2' and 3'

carbons of the carbohydrate. The catalyst is regenerated by reductive elimination; a base added to the reaction mixture removes the 3'-proton and halide to give the Pd (0) species again. The catalyst then can be reactivated and reentered into the coupling cycle.

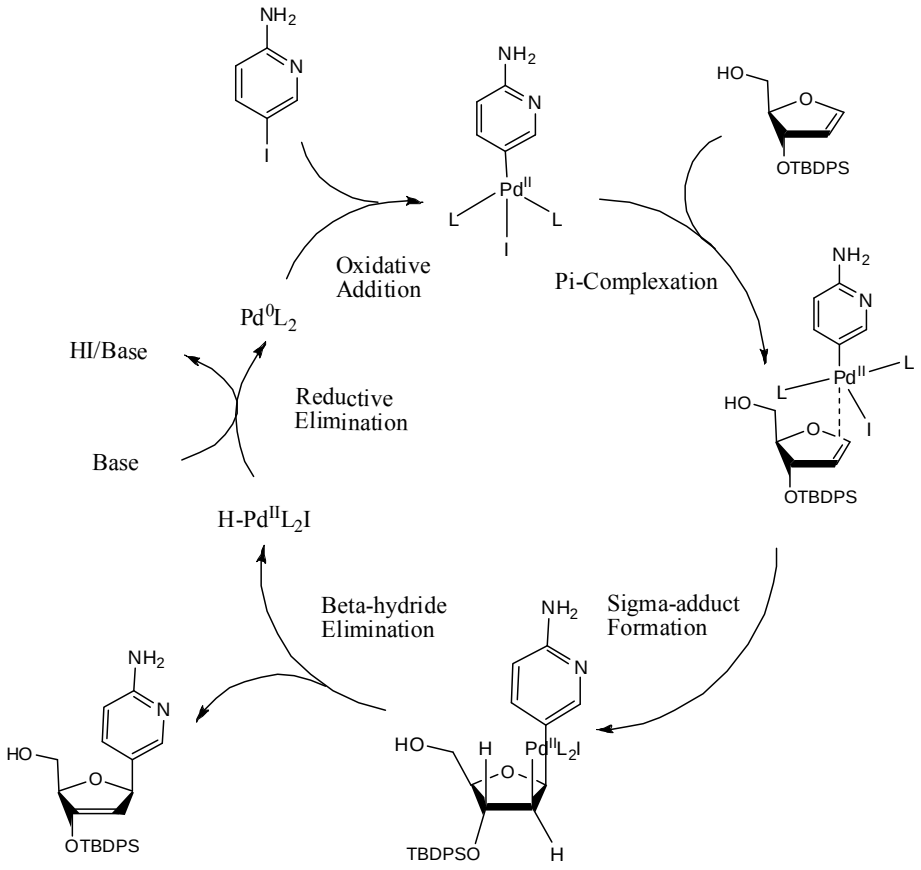


Figure 13. Mechanism of Heck coupling reaction.

The 2-aminopyridine derivative **6** had been synthesized successfully in our laboratory.[4] The results were reproducible and gave the desired product in good yield. For this reaction, the iodo-heterocycle was commercially available and so was sometimes used in excess. The proportion of the sugar to heterocycle was either 1 to 1 or 1 to1.5; extra heterocycle did not seem to help drive the reaction to completion. The catalyst

bis(dba)palladium(0) or tris(dba)dipalladium (0) made no difference in reaction condition or results (dba is a ligand: dibenzylideneacetone). When the reaction mixture was kept anhydrous and oxygen-free, the yield reached above 80 % easily.

Coupling reactions for the other *C*-nucleosides 7 – 10 seemed more sensitive to changes in reaction conditions. Deviations from the original procedures published by Dr. Zhenhua Sun[3] revealed more findings about this reaction, although yield was not necessarily enhanced. When the coupling reaction was first attempted for the 2-aminopyridine derivative **6**, 10 % pentafluorotriphenylphosphine ligand was first used for 10 % of tris(dba)dipalladium (0) in the reaction mixture. This resulted in very poor yield because the activated palladium (II) species required two neutral ligands, both of which may be phosphine. That is, 10 % of pentafluorotriphenylphosphine provided only 2.5 % of activated Pd (II) species after oxidative insertion. The yield was bound to be unappreciable due to the reduced amount of catalyst. Therefore, at least 20 % of the phosphine ligand was required for 10 % of the palladium catalyst.

Unlike the commercially available 2-amino-5-iodo-pyridine, all the other heterocycles were synthesized in the laboratory. The iodo-aromatic compounds were more precious compared to the carbohydrate, which can be produced in bulk for a large number of reactions. The heterocycle was therefore chosen to be the limiting reagent in the coupling reactions for **7** - **10** and the decision has proven wise. A general trend observed for these coupling reactions was that when the sugar was made in excess rather than the limiting reagent, the yield was also enhanced probably for a couple of reasons. First, this increases the chance of the sugar reacting with a palladium-inserted heterocycle complex at the beginning of the coupling cycle. Second, since the sugar decomposes at

elevated temperature over a prolonged period of time as this reaction requires, the rate of the reaction is made more competitive with the rate of decomposition by guaranteeing a surplus of that sugar.

Another observation was that higher reaction temperature dramatically improved yield. When the reaction mixture for synthesizing **8** was prepared as reported, refluxing at 55°C yielded only 3.6 % of the desired product. When the temperature was increased to 85°C, however, the yield was increased by about 15 fold (59 %). Higher temperature was probably increased the rate of the β-hydride elimination step.

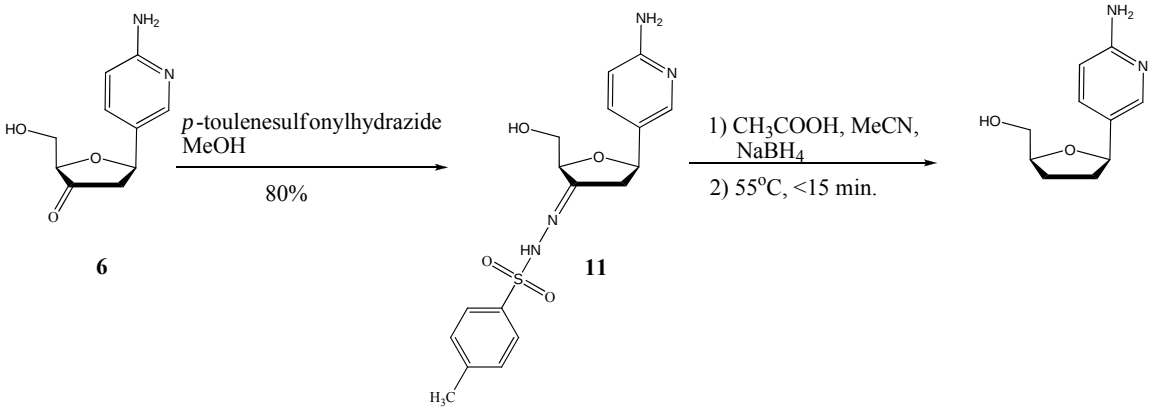
Removal of the 3'-O-TBDPS group of the direct products from the coupling reactions provided nucleosides **6** -**10**. The 2-aminopyridine nucleoside **6** was selected to convert into a dideoxy derivative. Two methods were attempted and both of them provided mild alternatives to standard Wolf-Kishner deoxygenation.[5] The first involves a tosylhydrazone intermediate and the second a dithioketal intermediate. Both employ reductive conditions to reduce the carbonyl of the sugar moiety into a methylene group.

Preparation of 2',3'-dideoxy-2-amino-pyridine (dd2APy)

The first route was reported by a former graduate student, Dr. Dongli Chen, in our laboratory (Scheme 3).[4a] He first converted **6** into a tosylhydrazone derivative **11** then used sodium triacetoxyborohydride [$Na(OAc)_3BH$] to reduce the hydrazone into the dideoxy nucleoside by stirring at ambient temperature for one hour. Six equivalents of sodium borohydride ($NaBH_4$) instead of a larger equivalent of the less reactive $Na(OAc)_3BH$ were used to completely consume the starting material. In fact, $NaBH_4$ in acetonitrile and acetic acid generates $Na(OAc)_3BH$ *in situ*[6]so either hydride source could

be used. The immediate product of this reduction step was *not* the desired dideoxy product. The R_f value of that compound was apparently distinct from that of the starting material. It was determined by ^{1}H NMR and mass spectral analysis to be a tosylhydrazide resulting from the direct reduction of the C-N double bond. Since the reported procedure did not give the final product, what more needed to be done to further reduce the hydrazide?

Scheme 3



A few literature references provided some mechanistic details.[5a,7] The reduction reaction proceeds in two stages: reduction to an intermediate which is then decomposed (Scheme 3). A tosylhydrazone is first reduced into a tosylhydrazide by hydrides such as $NaBH_4$ (Figure 14 A). Little hydrocarbon product is detected until the mixture is heated under reflux with water to yield a diazene intermediate, which then decomposed. Both the tosylhydrazide and diazene intermediates were detected using mass spectral analysis (Obs. 379.1430 for tosylhydrazide and 223.1205 for diazene) after heating for a short time. Two mechanisms have been proposed to account for the diazene decomposition.[8] The first is a chain reaction involving traces of oxygen or other oxidizing impurities

(Figure 14 B). A radical is first generated at the terminal nitrogen. Upon releasing a nitrogen gas molecule, an alkyl radical is generated, which then can be quenched by a hydrogen radical donor. The second mechanism requires the presence of a base (Figure 14 C). The base reversibly deprotonates the terminal nitrogen to give a diazenyl anion. Formation of a nitrogen gas molecule results in an alkyl anion. A proton transfer offers the saturated product. In either mechanism, thermal decomposition produces N_2, *p*-toluenesulphinic acid and alkane. In addition, the two hydrogen atoms of the methylene originate one from the reducing hydride and the second from the NH of the tosylhydrazone or from the water during workup.

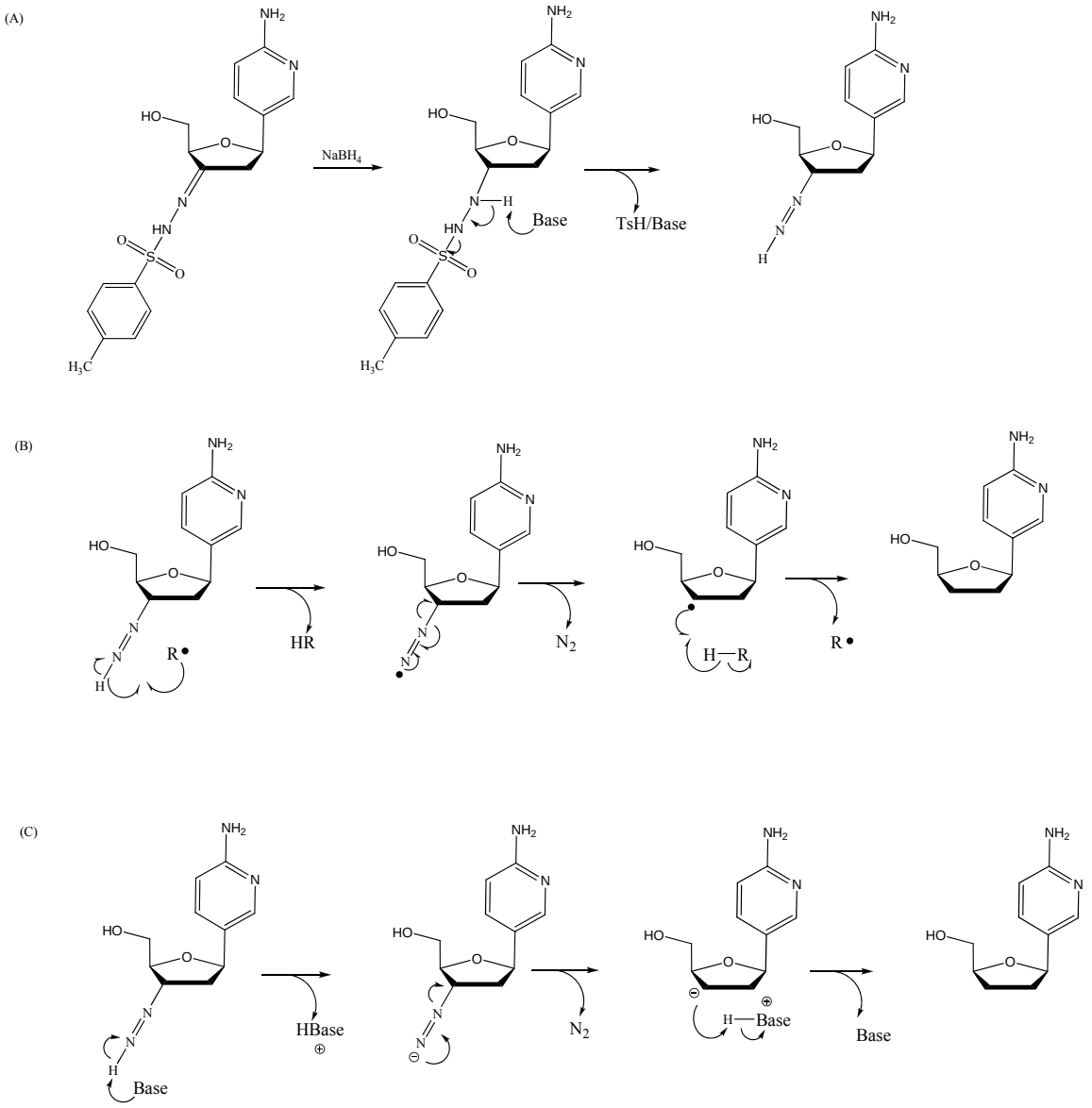
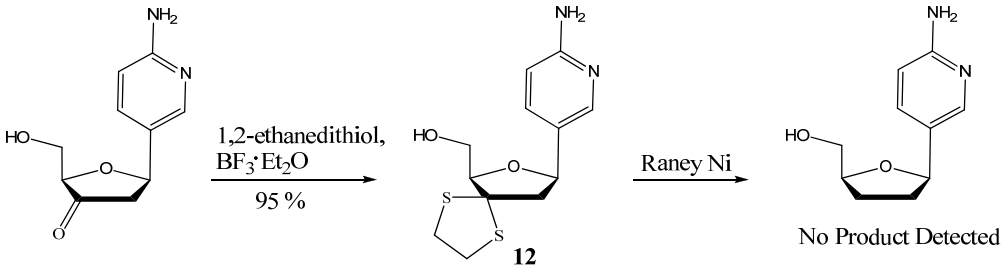


Figure 14. Mechanism of reduction and decomposition of tosylhydrazone to give dd2Py nucleoside. (A) Formation of the diazene intermediate. (B) and (C) are two proposed mechanism for the decomposition of the diazene to dideoxynucleosides dd2APy.

The right amount of heat was crucial for the reduction of the tosylhydrazide nucleoside. Without isolating the tosylhydrazide intermediate, solvents were removed on high vacuum to give an oil or a foam. Heating of that residue in ethanol at 65°C for an hour did not give the desired product. Further heating failed to show any change. It was later found that the tosylhydrazide intermediate had decomposed in the hot water bath of

the rotorvap! At elevated temperature, side reactions quenching the radical or anionic intermediates must have resulted in undesirable products, so subsequent heating of these unknown compounds failed to offer the desired dideoxy nucleoside. The temperature of the water bath for removing solvents must be kept at room temperature in order to preserve the intermediate and the decomposition conditions must be carefully monitored. The optimal temperature for the reaction was between 55 and 60°C. 15 minutes were all it took for the reaction to go to completion; prolonged heating would result in decomposition of the dideoxy nucleoside. A 1:1 mixture of ethanol and water was used to provide the hydrogen source during workup. Formation of nitrogen gas bubbles in the solution was always a positive sign that the reaction was working.

Scheme 4



The other method attempted for the synthesis of dd2APy involves a dithioacetal intermediate **12** (Scheme 4). The dithioacetal compound was readily produced in high yield in the presence of a catalytic amount of trifluoroboron etherate. [9] The boron Lewis acid chelates to the carbonyl of the sugar moiety and thereby increases the electrophilicity of the carbonyl carbon. 1,2-ethanedithiol then masks the carbonyl as a dithioacetal functional group.

Subsequent reduction by Raney nickel is also known as a desulfurization process as sulfur is removed from the dithioacetal compound. It is often done with a large excess

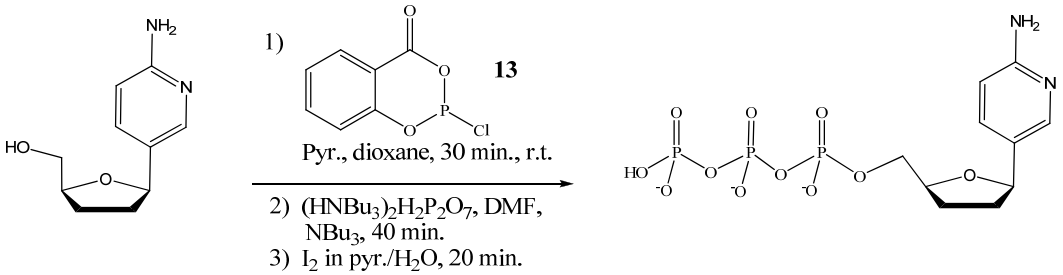
of Raney nickel (20 equivalents) in refluxing ethanol over a long period of time (at least 24 hours).[10] This method may be inappropriate for the synthesis of the dideoxynucleoside because nucleosides tend to be unstable at higher temperature especially when heated for a prolonged period of time. In addition, the use of hydrogen gas is not always required, although the solvent may provide a hydrogen source. The use of hydrogen gas in this reaction did not seem to help the process and a volume of relatively less polar solvent such as ethyl acetate was added to aid the solubility of the substrate. The Raney nickel generally used in the desulfurization reactions is freshly prepared as a W2 type.[11] The procedure involves repetitive alternate washing of aqueous NaOH and water upon a nickel-aluminum alloy. Since Raney nickel is pyrophoric when dry and exposed to air, a commercially available Raney nickel product numbered 2800 was tested for this reaction for safety and convenience reasons.[12] According to the Sigma-Aldrich catalogue, however, this product is suitable as a hydrogenation catalyst for aromatic, olefin, nitrile and nitro group reductions. This might be the main factor contributing to the failure of this reduction reaction as the activity and hydrogen content may not be appropriate for this desulfurization reaction. These difficulties hampered the development of this new method to synthesize the dideoxynucleoside product and the method was therefore abandoned.

Preparation and Purification of Triphosphates

The dd2APy nucleoside and two products from the Heck coupling reactions with bases of FAPy and MAPy were converted into 5'-triphosphate derivatives. There are many approaches to synthesizing nucleoside triphosphates but not one universal method

to prepare all modified derivatives.[13] Two synthetic procedures were compared and they differed only in the phosphorus activating reagent and workup. Ludwig and Eckstein developed a phosphite activating reagent **13**[14] which required an oxidation step (step 3) in the workup (Scheme 5). A less reactive phosphate activating reagent ($POCl_3$) is also widely used (Scheme 6).[15] Under either reaction condition, only the 5' hydroxyl group is reactive toward the activating compound. Addition of pyrophosphate gives a temporary cyclic intermediate, which is then hydrolyzed to give the corresponding 5'-triphosphate derivative.

Scheme 5



The Ludwig-Eckstein method employs a highly reactive and hydroscopic phosphorus (III) activator **13**. All solvents used in the method must be distilled one day ahead and stored in molecular sieves to ensure dryness. When performed carefully, the reaction could result in complete conversion to give the desired 5'-triphosphate derivative as indicated by a peak corresponding to product on an anion exchange column (Figure 15). The dd2APy triphosphate was synthesized successfully. However, compounds **7** and **8** were converted to monophosphates rather than triphosphates (data not shown). This suggests that the activated nucleoside might have been hydrolyzed prematurely due to the presence of water. Since the activating compound is a solid and taken out of the same

opened bottle for every use, water might have been introduced along with the addition of the reagent.

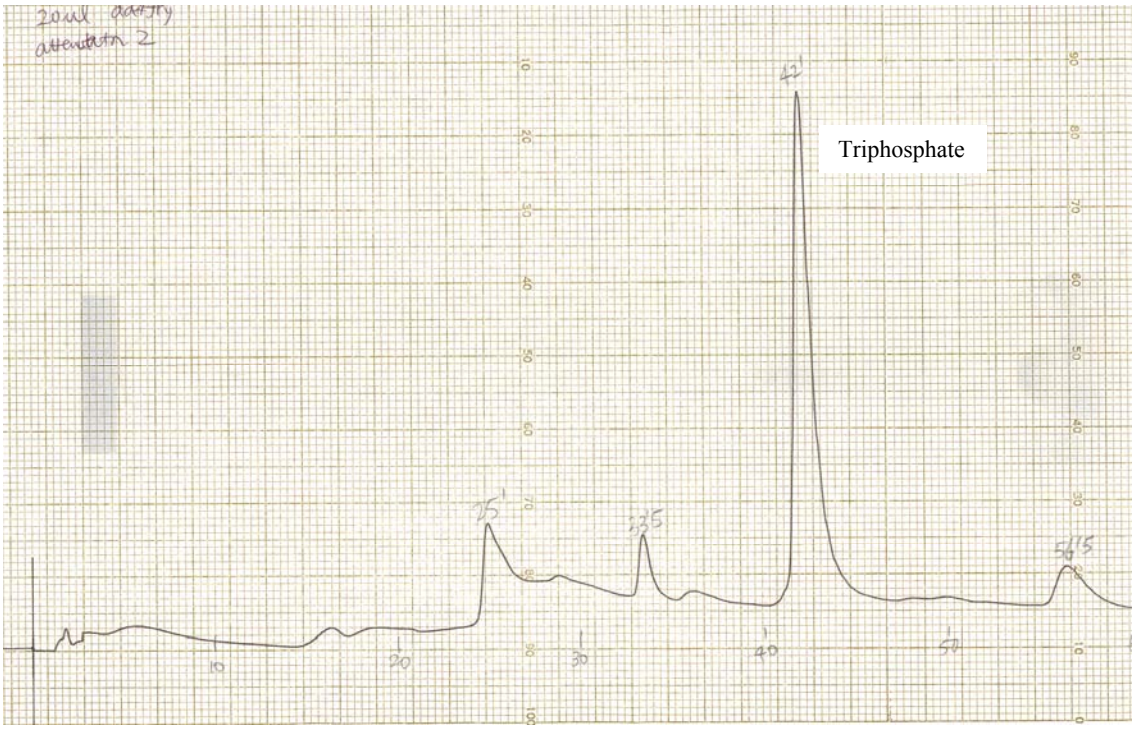
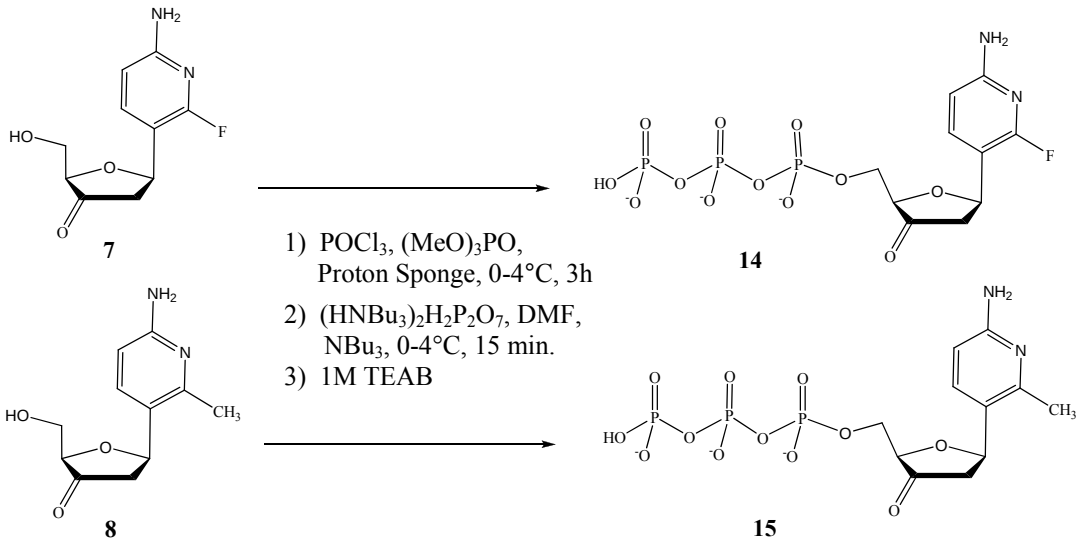


Figure 15. HPLC chromatogram of the crude mixture of dd2APyTP from an anion exchange column.

Scheme 6



In order to avoid untimely hydrolysis, we opted to select a less reactive phosphorus (V) reagent, $POCl_3$, for compounds **7** and **8**. Since the reagent is a liquid, it was possible to distill it and seal it until ready to use. This minimizes the exposure of the activator to water and air. The downside of this approach is reduced reactivity of phosphorus (V). The nucleoside, though activated, reacts more slowly with pyrophosphate than the equivalent activated nucleoside in the Ludwig-Eckstein method. This results in incomplete conversion into triphosphate, as indicated by the presence of an almost 1:1 ratio of mono- versus triphosphate (Figure 16).

Figure 16. HPLC chromatogram of the crude mixture of **15** from an anion exchange column.

All three products were purified on HPLC first by an anion exchange column followed by a reverse phase column. They were characterized by ^{31}P NMR (Figure 17)

and all had a signature peak near -22 ppm for the β-phosphorus atom, while the other two signals for α and γ phosphorus atoms occurred between -6 and -11 ppm depending on the pH and counterion.[13]

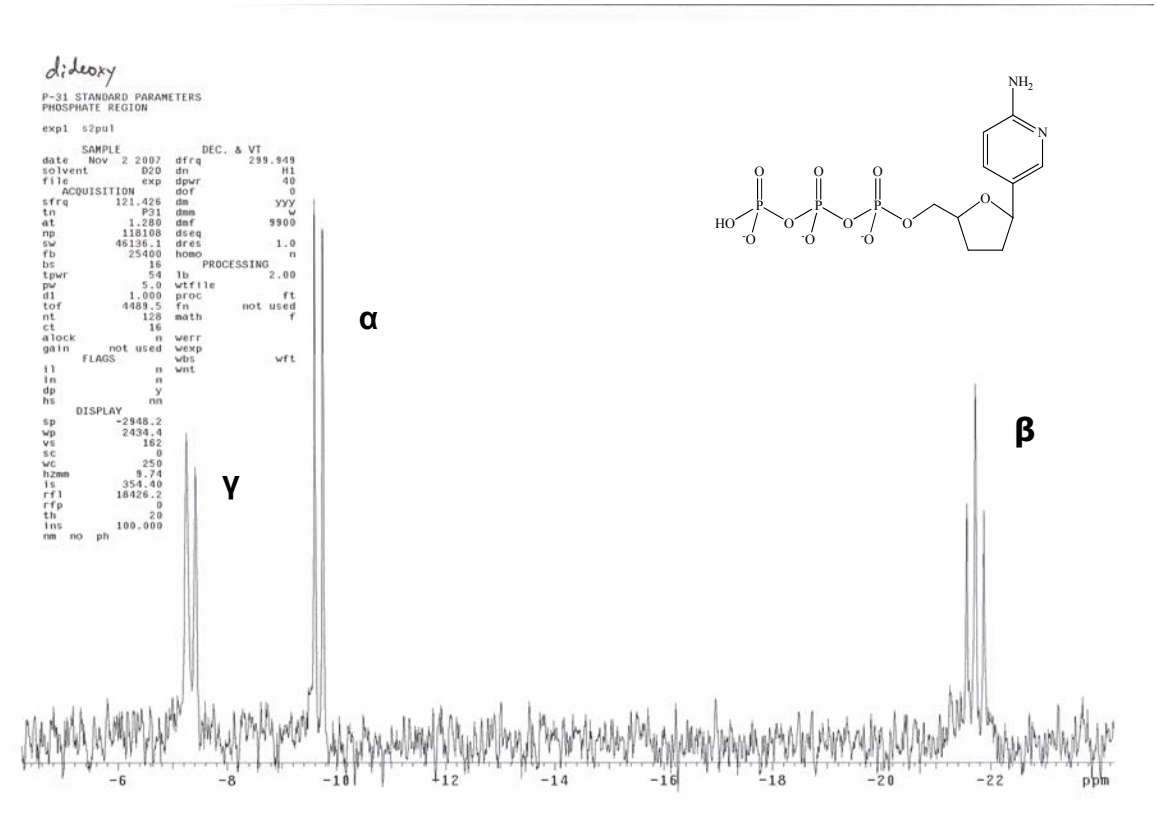


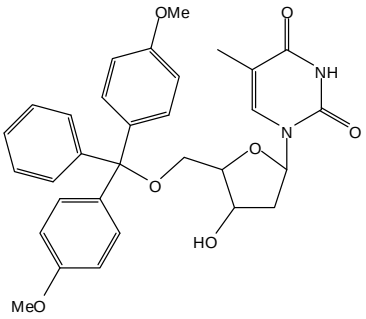
Figure 17. Sample ^{31}P NMR (300 MHz) of a nucleoside triphosphate.

Future Plans

One of the ultimate goals of this project is to investigate the role of minor groove interaction between the modified triphosphate substrate and the polymerase. The purified triphosphate derivatives will be tested in the future in various experiments, for example, enzyme-mediated DNA polymerization studies involving Pol α, Pol β, Pol γ, HIV-RT, and possibly other retroviruses. These studies will provide some insights into the nature and role of these minor groove interactions. What is most important – shape, size,

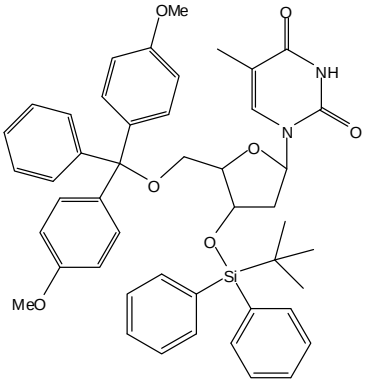
electronic effect, or hydrogen bonding requirement? If differential polymerase activity is observed for any of these triphosphate substrates, antiviral potentials will be examined.

Experimental Section



5'-O-(4,4'-Dimethoxytrityl)thymidine (2)

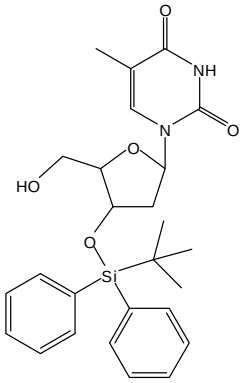
Thymidine (4 g, 16.5 mmol) was mostly dissolved in 40 mL of pyridine after two co-evaporations with that solvent. 4,4'-Dimethoxytrityl chloride (DMTr-Cl, 6.15 g, 18.2 mmol) was added, along with 4-(dimethylamino)pyridine (DMAP, 0.10 g, 0.825 mmol) and triethylamine (3.21 mL, 23.1 mmol). The homogeneous solution was stirred at ambient temperature for 2 hours. After removing the solvents, the mixture was redissolved in ethyl ether and water for extraction. The water removed unreacted thymidine and DMAP, while all DMT-containing compounds remained in the ethyl ether layer. The organic layer was dried with sodium sulfate, concentrated, and purified on a silica gel column. The product was eluted with 3 % MeOH/CH_2Cl_2 in the presence of 0.5 % triethylamine. The product was a white foam (90%, 8g). ^{1}H NMR was consistent with that previously reported.[3]



3'-O-(*t*-Butyldiphenylsilyl)-5'-O-(4,4'-Dimethoxytrityl)-thymidine (3)

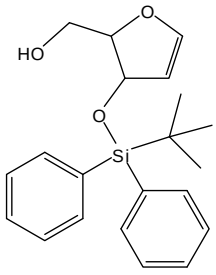
5'-O-DMT-thymidine (8 g, 14.7 mmol) and imidazole (2.5 g, 36.7 mmol) were dissolved in 30 mL of DMF. *t*-Butyldiphenylsilyl chloride (TBDPS-Cl, 4.2 mL, 16.4 mmol) was added dropwise at lowered temperature. The reaction was complete overnight. The solvent was removed and dichloromethane and brine were added for extraction. The

organic layer was dried and concentrated to an appropriate volume for the next reaction without purification.



3'-O-(*t*-Butyldiphenylsilyl)thymidine (4)

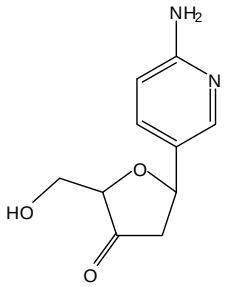
5'-O-DMT-3'-O-TBDPS-thymidine (5 g, 6.4 mmol) was dissolved in a mixture of 60 mL dichloromethane and about 15 mL methanol and cooled in an ice bath. 7 to 9 g of *p*-toluenesulfonic acid were added to drive the reaction to completion over the course of up to 1 hour. The acid was removed by washing with 5 % Na_2CO_3 solution twice. The organic layer was dried and concentrated to give a thick oil. The product was eluted with 40 % ethyl acetate in hexanes. It was a white foam (3g, 95 %). ^{1}H NMR was consistent with that previously reported. [3]



1,2-Dehydro-3-O-(*t*-butyldiphenylsilyl)-5-hydroxymethyl-furan (5)

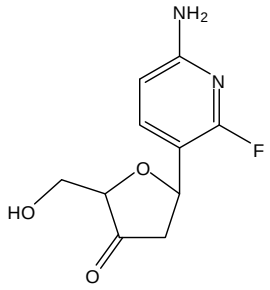
3'-O-TBDPS-thymidine (3.5 g, 7.3 mmol) was combined with ammonium sulfate (0.26 g, 2.0 mmol). 30 mL of 1,1,1,3,3,3-hexamethyldisilazane (HMDS) were injected into the refluxing apparatus. The mixture was heated at 140°C for 2 hours. Solvent was removed to give a thick cloudy mixture. 50 mL of methanol were added to produce a white suspension. At 0°C, 1.1 g of K_2CO_3 was added. After 30 minutes, the mixture was filtered through a thin layer of Celite. Solvents were removed to give a thick cloudy mixture again, which was immediately purified by silica gel with straight CH_2Cl_2. The

product was a clear oil (80 %, 2.8 g). ^{1}H NMR was consistent with that previously reported. [3]



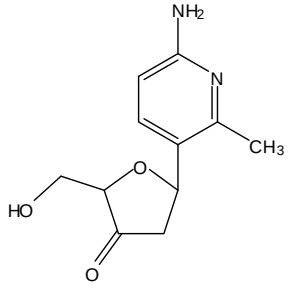
5(β-D-glyceropentofuran-3'-ulos-1')-2-amino-pyridine (6)

A mixture of tris(dibenzylideneacetone)dipalladium(0) (232 mg, 0.25 mmol) and pentafluorotriphenylphosphine (270 mg, 0.57 mmol) was stirred in 50 mL of acetonitrile at ambient temperature for 5 hours. It turned orange brown. Then a solution of the sugar **5** (1.0 g, 2.8 mmol) in 15 mL of acetonitrile from the previous step was injected into the catalyst mixture. 2-amino-5-iodo-pyridine (838 mg, 3.8 mmol) was added as a solid. N,N-diisopropylethylamine (126 µl, 0.72 mmol) was injected. After 48 hours of refluxing at 80°C, when the sugar was mostly consumed, the mixture was filtered through a thin layer of Celite. The brown filtrate was concentrated to a brown oil. The oil was dissolved in 5 mL of THF and cooled in an ice bath. Acetic acid (0.05 mL, 0.87 mmol) and 0.7 mL (2.4 mmol) of 1M tetrabutylammonium fluoride in THF were added. The brown solution was stirred on ice for 30 minutes. Solvents were removed and the product was eluted with 5 % MeOH/CH_2Cl_2 (1.0 g, 85 %). ^{1}H NMR was consistent with that previously reported.[2]



6-Amino-3-(β-D-glycero-pentofuran-3'-ulos-1'yl)-2-fluoropyridine (7)

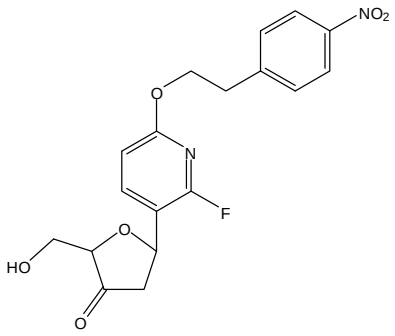
A mixture of tris(dibenzylideneacetone)dipalladium(0) (348 mg, 0.38 mmol) and pentafluorotriphenylphosphine (810 mg, 1.5 mmol) was stirred in 50 mL of acetonitrile at ambient temperature for 5 hours. It turned orange brown. Then a solution of the sugar **5** (672 mg, 1.9 mmol) in 15 mL of acetonitrile from the previous step was injected into the catalyst mixture. 6-amino-2-fluoro-3-iodopyridine (300 mg, 2.3 mmol) was added as a solid. N,N-diisopropylethylamine (198 μl, 1.1 mmol) was injected. After 48 hours of refluxing at 85°C, when the heterocycle was mostly consumed, the mixture was filtered through a thin layer of Celite. The brown filtrate was concentrated to a brown oil. The oil was dissolved in 5 mL of THF and cooled in an ice bath. Acetic acid (0.05 mL, 0.87 mmol) and 0.3 mL (1.1 mmol) of 1M tetrabutylammonium fluoride in THF were added. The brown solution was stirred on ice for 30 minutes. Solvents were removed and the product was eluted with 1.5 % MeOH/CH_2Cl_2 (133 mg, 59 %). ^{1}H NMR was consistent with that previously reported.[2]



6-Amino-3-(β-D-glycero-pentofuran-3'-ulos-1'yl)-2-methylpyridine (8)

A mixture of bis(dibenzylideneacetone)palladium(0) (128 mg, 0.22 mmol) and pentafluorotriphenylphosphine (426 mg, 0.80 mmol) was stirred in 50 mL of acetonitrile at ambient temperature for 5 hours. It turned orange brown. Then a solution of the sugar **5** (830 mg, 2.3 mmol) in 15 mL of acetonitrile from the previous step was injected into the catalyst

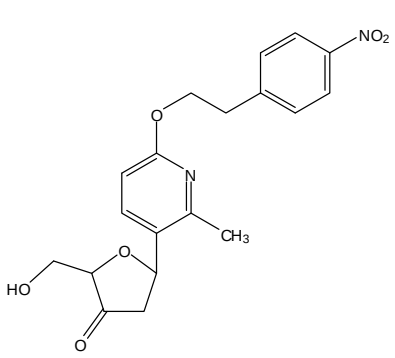
mixture. 6-amino-2-methyl-3-iodopyridine (262 mg, 2.1 mmol) was added as a solid. N,N-diisopropylethylamine (389 μl, 2.2 mmol) was injected. After 48 hours of refluxing at 90-95°C, when the heterocycle was mostly consumed, the mixture was filtered through a thin layer of Celite. The brown filtrate was concentrated to a brown oil. The oil was dissolved in 5 mL of THF and cooled in an ice bath. Acetic acid (0.05 mL, 0.87 mmol) and 0.3 mL (1.1 mmol) of 1M tetrabutylammonium fluoride in THF were added. The brown solution was stirred on ice for 30 minutes. Solvents were removed and the product was eluted with 5 % MeOH/CH_2Cl_2 (150 mg, 29 %). ^{1}H NMR was consistent with that previously reported.[2]



6-[2-(4-Nitrophenlethoxy)]-3-(β-D-glyceropentofuran-3'-ulos-1'yl)-2-fluropyridine (9)

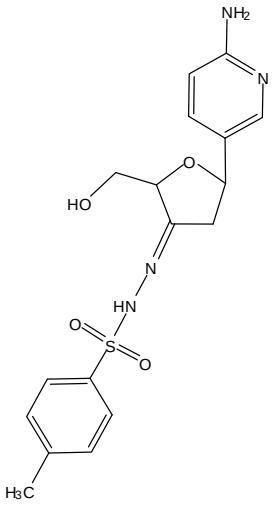
A mixture of bis(dibenzylideneacetone)palladium(0) (128 mg, 0.22 mmol) and pentafluorotriphenyl-phosphine (426 mg, 0.80 mmol) was stirred in 50 mL of acetonitrile at ambient temperature for 5 hours. It turned orange brown. Then a solution of the sugar **5** (830 mg, 2.3 mmol) in 15 mL of acetonitrile from the previous step was injected into the catalyst mixture. 6-(4-nitrophenylethoxy)-2-fluoro-3-iodopyridine (433 mg, 1.1 mmol) was added as a solid. N,N-diisopropylethylamine (389 μl, 2.2 mmol) was injected. After 48 hours of refluxing at 90°C, when the heterocycle was mostly consumed, the mixture was filtered through a thin layer of Celite. The brown filtrate was concentrated to a brown oil. The oil was dissolved in 5 mL of THF and cooled in an ice bath. Acetic acid (0.05 mL, 0.87 mmol)

and 0.3 mL (1.1 mmol) of 1M tetrabutylammonium fluoride in THF were added. The brown solution was stirred on ice for 30 minutes. Solvents were removed and the product was eluted with 0.5 % MeOH/CH_2Cl_2 (200 mg, 23 %). ^{1}H NMR was consistent with that previously reported.[2]



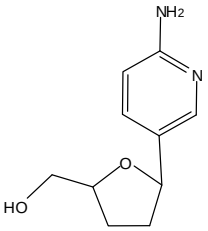
6-[2-(4-Nitrophenlethoxy)]-3-(β-D-glyceropentofuran-3'-ulos-1'yl)-2-methylpyridine (10)

A mixture of tris(dibenzylideneacetone)dipalladium(0) (150 mg, 0.16 mmol) and pentafluorotriphenyl-phosphine (120 mg, 0.23 mmol) was stirred in 50 mL of acetonitrile at ambient temperature for 5 hours. It turned orange brown. Then a solution of the sugar **5** (500 mg, 1.4 mmol) in 15 mL of acetonitrile from the previous step was injected into the catalyst mixture. 6-(4-nitrophenylethoxy)-3-iodo-2-methylpyridine (800 mg, 2.1 mmol) was added as a solid. N,N-diisopropylethylamine (100 µl, 0.57 mmol) was injected. After 48 hours of refluxing at 90°C, when the heterocycle was mostly consumed, the mixture was filtered through a thin layer of Celite. The brown filtrate was concentrated to a brown oil. The oil was dissolved in 5 mL of THF and cooled in an ice bath. Acetic acid (0.05 mL, 0.87 mmol) and 0.3 mL (1.1 mmol) of 1M tetrabutylammonium fluoride in THF were added. The brown solution was stirred on ice for 30 minutes. Solvents were removed and the product was eluted with 0.5 % MeOH/CH_2Cl_2 (200 mg, 23 %). ^{1}H NMR was consistent with that previously reported.[2]



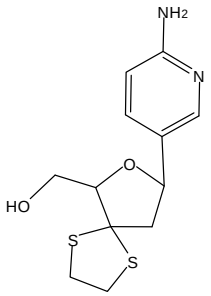
5-(β-D-glyceropentofuran-3'-ulos-1'-yl)-2-aminopyridine p-toluenesulfonylhydrazone (11)

Compound **6** (200 mg, 0.96 mmol) was suspended in about 4 mL of freshly distilled methanol. *p*-Toluenesulfonylhydrazide (0.54 g, mmol) was added. The mixture was stirred at ambient temperature overnight. The reaction was monitored by alumina TLC plate. Rf :0.2 (5 % methanol in dichloromethane). The product precipitate was collected and the filtrate was subjected to silica gel column. More product was eluted with 5 % MeOH/CH_2Cl_2 (total 300 mg, 83 %). ^{1}H NMR was consistent with that previously reported.[4a]



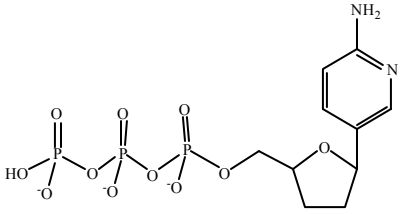
5(β-D-glyceropentofuran-2',3'-dideoxy-1')-2-amino-pyridine (dd2APy)

The hydrazone (300 mg, mmol) was suspended in 4 mL of dry acetonitrile. About 3 mL of acetic acid were added dropwise at 0°C. The slurry was stirred to gradually turn clear. $NaBH_4$ (60 mg, mmol) was added in 2 portions over a 45 min. period. The reduction step was over as indicated by the disappearance of starting material and the appearance of a slower running spot. About 1-2 mL of water was used to quench the reaction; bubbles evolved and stopped. Solvents were evaporated at ambient temperature, using ethanol to coevaporate the acid. The resulted yellow oil was redissolved in 3 mL of a 1:1 mixture of ethanol and water. The solution was heated at 55-60°C for 15 min. to give a darker yellow solution. The product was eluted with 5% MeOH/CH_2Cl_2. ^{1}H NMR was consistent with that previously reported.[4a]



(8-(6-aminopyridin-3-yl)-7-oxa-1,4-dithiaspiro[4.4]nonan-6-yl)methanol (12)

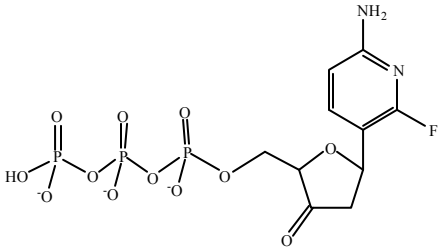
Compound **6** (230 mg, 1.1 mmol) was dissolved in 15 mL of CH_2Cl_2. the yellow solution was cooled to -78°C. $BF_3 \cdot Et_2O$ (0.68 mL, 5.55 mmol) was added dropwise. 1,2-ethanedithiol (3.3 mL, 2.8 mmol) was added to give an orange mixture. The mixture was stirred at -78°C for 30 minutes then at ambient temperature overnight. The solution turned clear and light yellow. Solvents including1,2-ethanedithiol was removed on vacuum. The product was eluted with 2 % MeOH/CH_2Cl_2 with 0.5 % TEA. The product was a solid (300 mg, 96%). Rf (7 % MeOH/CH_2Cl_2) = 0.31. ^{1}H NMR ($CDCl_3$) δ = 2.45 (dd, 1H), 2.75 (dd, 1H), 3.34 (m, 2H), 3.38 (m, 2H), 3.85 (d, 2H), 4.20 (m, 1H), 4.45 (b, 2H), 4.90 (m, 1H), 6.50 (d, 1H), 7.50 (d, 1H), 8.03 (s, 1H). ^{13}C NMR ($CD_3OD/CHCl_3$) δ = 39.8, 41.1, 54.7, 64.6, 71.1, 78.8, 79.5, 90.1, 110, 127, 139, 146. Exact mass calculated for [$C_{12}H_{16}N_2O_2S_2$] requires 284.07; found, 284.6227 (ESI^+).



5'-Triphosphate derivative of dd2APy (14)

The dideoxynucleosides dd2APy (15 mg, 0.077 mmol) was coevaporated with pyridine twice and dried in a desiccator for at least 1 h. 77 µl of pyridine and 231 µl of freshly distilled dioxane were added to dissolve the starting material. 77 µl of a 1M solution of 2-Cl-benzodioxaphosphorin-4-one in dioxane were added. Precipitation occurred as the pyridine precipitated with chloride. 30 min. later, 231 µl of a 0.5 M solution of tetrabutylammonium pyrophosphate in DMF were injected. The mixture turned very cloudy until 77 µl of dry tributylamine was added. 40 min. later, 2 mL of 1% I_2 in 98:2

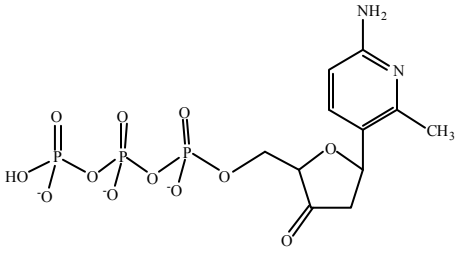
pyridine:water mixture were added for oxidation. After 20 min., solvents were removed on high vacuum to give an oil. Ether and water extractions resulted in yellow layers. The aqueous one was concentrated and filtered for HPLC purification. The mixture was first separated on an anion-exchange column to isolate the triphosphate using a 30 min. gradient of 0 – 100% Buffer B. Buffer A was deionized water and Buffer B was 1M triethylammonium bicarbonate (TEAB). After concentration of the fractions, the sample was desalted on a reverse phase column using a 15 min. gradient of 0 – 100 % Buffer B. Buffer A was 50 mM triethylammonium acetate (TEAA) pH 7.0 in 2 % methanol. Buffer B contained the same salt concentration with 50 % methanol. The product was concentrated and coevaporated with water a few times on the lyophilizer. λ_{max} (H_2O) = 290 nm. ^{31}P NMR (300 MHz, D_2O) δ = -9 (d, 1P), -7 (d, 1P), -22 (t, 1P).



5'-Triphosphate derivative of 7 (15)

The nucleoside **7** (18 mg, 0.080 mmol) was coevaporated with acetonitrile twice to give a yellow foam and dried in a desiccator overnight. It was dissolved in trimethylphosphate (380 µl) at 0-4°C. 10% proton sponge (2 mg) and freshly distilled $POCl_3$ (18.2 µl, 0.080 mmol) were added at once. The brown solution turned slightly darker in color. 3 h. later, all starting material was consumed. A solution of bis-tributylammonium pyrophosphate (0.5 M in DMF, 0.86 mL, 0.43 mmol) and tributylamine (200 µl, 0.84 mmol) were added quickly. The reaction was quenched 15 min. later with 1 M triethylammonium bicarbonate buffer (TEAB, pH 7.5, about 6 mL). Solvents were removed, using ethanol for a few coevaporations. The resulted solid was

redissolved in water. The yellow solution was filtered and purified on HPLC. The purification process on the ion exchange column was the same as that described for **14**. The sample still remains to be desalted on a reverse phase column. λ_{max} (H_2O) = 290 nm. ^{31}P NMR (300 MHz, D_2O) of the pre-desalted sample δ = -10, -22 .



5'-Triphosphate derivative of 8 (16)

The nucleoside **8** (33 mg, 0.135 mmol) was coevaporated with acetonitrile twice to give a yellow foam and dried in a desiccator overnight. It was dissolved in trimethylphosphate (675 µl) at 0-4°C. 10% proton sponge (2 mg) and freshly distilled $POCl_3$ (31 µl, 0.35 mmol) were added at once. The brown solution turned slightly darker in color. 3 h. later, all starting material was consumed. A solution of bis-tributylammonium pyrophosphate (0.5 M in DMF, 1.3 mL, 0.65 mmol) and tributylamine (338 µl, 1.60 mmol) were added quickly. The reaction was quenched 15 min. later with 1 M triethylammonium bicarbonate buffer (TEAB, pH 7.5, about 10 mL). Solvents were removed, using ethanol for a few coevaporations. The resulted solid was redissolved in water. The yellow solution was filtered and purified on HPLC. The purification process on the ion exchange column was the same as that described for **14**. λ_{max} (H_2O) = 290 nm. ^{31}P NMR (300 MHz, D_2O) δ = -9.5 (1P, d), -10.5 (1P, d), -22 (1P, t).

References

1 (a) Crisp, G. T. *Chemical Society Reviews* **1998**, 27, 427-436. (b) Doyle, Jr. G. D. *Acc. Chem. Res.* **1990**, 23, 201-206.

2 Sun, Z., Ahmed, S. & McLaughlin, L. W. *J. Org. Chem.* **2006**, 71, 2922-2925.

3 Lan, T. & McLaughlin, L. W. *JACS* **2000**, 122, 6512-6513.

4 (a) Fraley, A. W., Chen, D., Johnson, K., McLaughlin, L. W. *JACS* **2003**, 125(3), 616-617. (b) Searls, T., Chen, D. L., Lan, T., McLaughlin, L. W. *Biochemistry* **2000**, 39(15) 4375-4382.

5 (a) Hutchins, R. O., Milewski, C. A. & Maryanoff, B. E. *JACS* **1973**, 95, 3662. (b) Hutchins, R. O., Kacher, M. & Rua, L. *J. Org. Chem*. **1975**, 40, 923. (c) Taylor, E. J. & Djerassi, C. *JACS* **1976**, 98, 2275.

6 Hutchins, R. O. & Natale, N. R. *J. Org. Chem.* **1978**, 43, 11, 2299-2231.

7 Caglioti, L. *Tetrahedron* **1996**, 22, 487-493.

8 Tsuji, T. & Kosower, E. M. *JACS* **1971**, 93, 8, 1992-1999.

9 Davis, F. A. & Santhanaraman, M. *J. Org. Chem.* **2006**, 71, 4222-4226.

10 (a) Wolfrom, M. L. & Karabinos, J. V. *JACS* **1944**, 66, 909-911. (b) Canet, J., Fadel, A. & Salun, J. *J. Org. Chem.* **1992**, 57, 3463-3473.

11 (a) Augustine, R. L. *Catalytic hydrogenation* Dekker, M., Ed., M. Dekker: New York, **1965**, 23. (b) Mozingo, R., Adkins, H. & Richards, L. *Organic Syntheses* **1941**, 21, 15-17.

12 Mebane, R. c. Jensen, D. R., Richerd, K. R. & Gross B. H. *Synthetic Communications* **2003**, 33, 19, 3373-3379.

13 Burgess, K. & Cook, D. *Chem. Rev.* **2000**, 100, 2047-2059.

14 Ludwig, J. & Eckstein, F. *J. Org. Chem.* **1989**, 54, 631-635.

15 Kovacs, T. Z. & Otvos, L. S. *Tetrahedron Lett.* **1988**, *29*, 4525.